

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 10, 2003, 13:51:54 ; Search time 21 Seconds
(Without alignments)
1716.686 Million cell updates/sec

Title: US-09-839-185-2

Perfect score: 2040
Sequence: 1 MEMSGNSGPGHGPQASGSG.....MEDENTRAYDSSSHHTNWSL 375

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2033	99.7	375	2 T00929	squamosa-promoter
2	1956	95.9	373	2 T52553	squamosa-promoter
3	863.5	42.3	354	2 T46014	squamosa-promoter
4	615	30.1	323	2 T52297	squamosa-promoter
5	556	27.3	257	2 T52298	squamosa-promoter
6	400	19.6	425	2 T52604	squamosa-promoter
7	392	19.2	419	2 T52603	squamosa-promoter
8	374.5	18.4	405	2 T52582	squamosa-promoter
9	356	17.5	396	2 T52586	squamosa-promoter
10	354.5	17.4	393	2 B86359	protein F17L21.14
11	333	16.3	333	2 T52598	squamosa-promoter
12	333	16.2	171	2 T52594	squamosa-promoter
13	330	16.2	181	2 T52607	squamosa-promoter
14	328.5	16.1	1020	2 H96793	unknown protein F1
15	321.5	15.8	138	2 T52567	squamosa-promoter
16	321	15.7	927	2 T47827	squamosa-promoter
17	318.5	15.6	1035	2 G86342	hypothetical prote
18	313	15.3	305	2 T52289	squamosa-promoter
19	310	15.2	881	2 T52601	squamosa-promoter
20	309	15.1	881	2 T52602	squamosa-promoter
21	308.5	15.0	131	2 T04328	liguleless1 protei
22	305	14.9	167	2 T52599	squamosa-promoter
23	303.5	14.9	174	2 T52600	squamosa-promoter
24	297	14.6	240	2 T02180	probable squamosa-
25	292.5	14.3	129	2 H84749	squamosa-promoter
26	292.5	14.3	131	2 T52597	squamosa-promoter
27	255.5	12.5	377	2 T52606	squamosa-promoter

30	254.5	12.5	801	2 T52605	squamosa promoter
31	160.5	7.9	1043	2 T13733	FT2-F1 protein - f
32	157.5	7.7	331	2 G96785	protein-F10A5.24 l
33	152	7.5	1584	2 T18276	protein-tyrosine k
34	146	7.2	1858	2 T18273	1-phosphatidylinos
35	145	7.1	704	2 S21911	BRcore-NS-23 prote
36	140	6.9	695	2 T21950	hypothetical prote
37	129	6.3	1555	2 T13998	gene mastermind pr
38	127	6.2	602	2 F84432	probable C2H2-type
39	126	6.2	1390	2 T14004	trfA protein - sil
40	126	6.2	1748	2 S42136	cnb protein - Tet
41	125	6.1	647	2 S05450	steroid hormone re
42	125	6.1	864	2 A49070	ecdysone-inducible
43	124.5	6.1	603	2 F96797	probable Gr-like t
44	124	6.1	1230	2 T18259	serine/threonine p
45	123.5	6.1	758	2 S45825	cell division cont

ALIGNMENTS

RESULT 1

T00929
squamosa-promoter binding protein 9 [imported] - Arabidopsis thaliana

N/Alternate names: transcription factor sp19

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 12-Feb-1999 #sequence revision 12-Feb-1999 #text_change 23-Mar-2001

C/Accession: T00929; T52595; A84851

R/Rounley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul

submitted to the EMBL Data Library, December 1997

A/Description: Arabidopsis thaliana chromosome II BAC T24P15 genomic sequence.

A/Reference number: Z14212

A/Accession: T00929

A/Status: translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-375 <ROU>

A/Cross-references: EMBL:AC00561; NID:g2673901; PID:g2673911

A/Experimental source: cultivar Columbia

R/Caron, G.H.; Hoehmann, S.; Klein, J.; Nettesheim, K.; Saedler, H.; Huijser, P.

Gene 237, 91-104, 1999

A/Title: Molecular characterisation of the Arabidopsis SBP-box genes.

A/Reference number: Z52336; WUID:99453765; PMID:10524240

A/Accession: T52595

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-375 <CAR>

A/Cross-references: EMBL:AO11638; PIDN:CAB56590.1

A/Experimental source: cultivar Columbia; tissue type inflorescence

R/Lin, X.; Kaul, S.; Rounley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanKen, S.E.; Umayam, L.; Tallon, L.

euse, D.; Nieman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A/Reference number: A84420; WUID:20083487; PMID:10617197

A/Accession: A84851

A/Status: preliminary

A/Molecule type: DNA

A/Cross-references: GB:AE002093; NID:g2673911; PIDN:AAB8645.1; GSPDB:GN00139

C/Genetics:

A/Gene: sp19; T24P15.11; AC2942200

A/Map position: 2

A/Intons: 120/2; 169/1

C/Keywords: DNA binding; transcription factor

Query Match 99.7%; Score 2033; DB 2; Length 375;

Best Local Similarity 99.7%; Pred. No. 6e-146;

Matches 374; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MEMSGNSGPGHGPQASGSGSTSSSFSGGLMFGCKITFEDGGSGSSGGRNRRV 60
Db 1 MEMSGNSGPGHGPQASGSGSTSSSFSGGLMFGCKITFEDGGSGSSGGRNRRV 60

Qy 61 RGGSGGSGQIPRCVCEGCGMDLTNAKGYSHRVCGVSHKTPKVTAGIEGRFCQCCSR 120
 |||||
 Db 61 RGGSGGSGQIPRCVCEGCGMDLTNAKGYSHRVCGVSHKTPKVTAGIEGRFCQCCSR 120
 Qy 121 FHQLEPFLERKSCRRRLAGHNERRRKPPASLSTLASRYGIAISLYENGDAQNGSFL 180
 |||||
 Db 121 FHQLEPFLERKSCRRRLAGHNERRRKPPASLSTLASRYGIAISLYENGDAQNGSFL 180
 Qy 181 GNOEIGWSSSRLDTRVMRPPVSSPSMOINPMNVPSQSGVGGRTSFSPEIMDTKLESY 240
 |||||
 Db 181 GNOEIGWSSSRLDTRVMRPPVSSPSMOINPMNVPSQSGVGGRTSFSPEIMDTKLESY 240
 Qy 241 KGIGDNCALSLISNPHQPHDNNNNNNNNNNNNNTWPASSGFGPMVTYMAQPPAPSQH 300
 |||||
 Db 241 KGIGDNCALSLISNPHQPHDNNNNNNNNNNNNNTWPASSGFGPMVTYMAQPPAPSQH 300
 Qy 301 QYLNPPWFKNDNDMSPLYNLGRTTEPDNCISSTGTAMGEFELSDHHQSRQYMEDN 360
 |||||
 Db 301 QYLNPPWFKNDNDMSPLYNLGRTTEPDNCISSTGTAMGEFELSDHHQSRQYMEDN 360
 Qy 361 TRAYDSSSHHTWMSL 375
 |||||
 Db 361 TRAYDSSSHHTWMSL 375

RESULT 2.

T52593

squamosa promoter binding protein homolog 9 (imported) - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear crease)

C:Date: 24-Oct-2000 #sequence_revision 24-Oct-2000 #text_change 24-Oct-2000

C:Accession: T52593

R:Cardon, G.H.; Hohmann, S.; Nettesheim, K.; Saedler, H.; Huijser, P.

Plant J. 12, 367-377, 1997

A:Title: Functional analysis of the Arabidopsis thaliana SBP-box gene SPL3: a novel gene

A:Reference number: Z24029; MUID:9301089; PMID:9301089

A:Accession: T52593

A:Status: Preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-373 <CAR>

A:Cross-references: EMBL:AJ011639; PIDN:CAB56591.1

A:Experimental source: cultivar Landsberg erecta; inflorescence

C:Genetics:

A:Gene: spl3

A:Map position:

C:Keywords: DNA binding; transcription factor

Query Match

Best Local Similarity

Matches 366; Conservative

0; Mismatches

1; Indels

14; Gaps

2;

Qy 1 MEMGNSGPHQPGQAESGSSSTESSFSGGLMFGQKIYFEDGGGSGSSSGGRNRRV 60
 |||||
 Db 1 MEMGNSGPHQPGQAESGSSSTESSFSGGLMFGQKIYFEDGGGSGSSSGGRNRRV 60
 Qy 61 RGGSGGSGQIPRCVCEGCGMDLTNAKGYSHRVCGVSHKTPKVTAGIEGRFCQCCSR 120
 |||||
 Db 61 RGGSGGSGQIPRCVCEGCGMDLTNAKGYSHRVCGVSHKTPKVTAGIEGRFCQCCSR 120
 Qy 121 FHQLEPFLERKSCRRRLAGHNERRRKPPASLSTLASRYGIAISLYENGDAQNGSFL 180
 |||||
 Db 121 FHQLEPFLERKSCRRRLAGHNERRRKPPASLSTLASRYGIAISLYENGDAQNGSFL 180
 Qy 181 GNOEIGWSSSRLDTRVMRPPVSSPSMOINPMNVPSQSGVGGRTSFSPEIMDTKLESY 240
 |||||
 Db 181 GNOEIGWSSSRLDTRVMRPPVSSPSMOINPMNVPSQSGVGGRTSFSPEIMDTKLESY 240
 Qy 241 KGIGDNCALSLISNPHQPHDNNNNNNNNNNNNNTWPASSGFGPMVTYMAQPP 294
 |||||
 Db 241 KGIGDNCALSLISNPHQPHDNNNNNNNNNNNNNTWPASSGFGPMVTYMAQPP 294
 Qy 295 PAPSOHOYLNPPWFKNDNDMSPLYNLGRTTEPDNCISSTGTAMGEFELSDHHQSRQ 354
 |||||
 Db 301 PAPSOHOYLNPPWFKNDNDMSPLYNLGRTTEPDNCISSTGTAMGEFELSDHHQSRQ 352

Qy 355 YMEDENTRAYDSSSHHTWMSL 375
 |||||
 Db 353 YMEDENTRAYDSSSHHTWMSL 373

RESULT 3

T46014

squamosa promoter-binding protein homolog - Arabidopsis thaliana

N:Alternate names: protein T10K17.130

C:Species: Arabidopsis thaliana (mouse-ear crease)

C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000

C:Accession: T46014

R:Benes, V.; Wurbach, E.; Drzonek, H.; Ansoerge, W.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.

submitted to the Protein Sequence Database, January 2000

A:Reference number: Z23019

A:Accession: T46014

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-354 <BEN>

A:Cross-references: EMBL:AL132977

A:Experimental source: cultivar Columbia; BAC clone T10K17

C:Genetics:

A:Map position: 3

A:Introns: 105/2; 153/1

A:Note: T10K17.130

Query Match

Best Local Similarity

Matches 204; Conservative

32; Mismatches

86; Indels

65; Gaps

17;

Qy 14 GQAESGG-SSTESSFSGGLMFGQKIYFEDGGGSGSSSGGRNRRVRRGGSGGQIP 72
 |||||
 Db 8 GQAESGSSSTESSSLGGLMFGQKIYFEDG-----SGSRKNRV--NTVKSSTTA 57
 Qy 73 RCQVGGCGMDLTNAKGYSHRVCGVSHKTPKVTAGIEGRFCQCCSRFHQLPFDLEKR 132
 |||||
 Db 58 RCQVGGCGMDLTNAKGYSHRVCGVSHKTPKVTAGIEGRFCQCCSRFHQLPFDLEKR 117
 Qy 133 SCRRLAGHNERRRKPPASLSTLASRYGIAISLYENGDAQNGSFLGNOEIMPSSRT 192
 |||||
 Db 118 SCRRLAGHNERRRKPPASLSTLASRYGIAISLYENGDAQNGSFLGNOEIMPSSRT 175
 Qy 193 LDRVMRPPVSSPSMOINP-----NMVPSQSGVGGRTSFS--PEIMDTKLESYKIG 244
 |||||
 Db 176 ---VMQRP---GPMQINPVARETHPHMNVLSHGS-----SSFTTCPEMINNST----- 217
 Qy 245 DSNCALSLISNPHQPHDNNNNNNNNNNNNNTWPASSGFGPM-----TYTMAQF----- 293
 |||||
 Db 218 DSNCALSLISNPHQPHDNNNNNNNNNNNNNTWPASSGFGPMISFSDKVTVAQPPPISTH 271
 Qy 294 -PPAPSOHOYLNPPWFKNDNDMSPLYNLGRTTEPDNCISSTGTAMGEFELSDHH 348
 |||||
 Db 212 QPPISTHOQVLSQTWEEVYAGKSHHYNSPV---SQISEPADPQISNGTMMGFEELY-LH 327
 Qy 349 HQSRQYMEDENTRAYDSSSHHTWMSL 375
 |||||
 Db 328 QQVLYKQVMEPENTRAYDSSPHFWMSL 354

RESULT 4

T52297

squamosa promoter binding protein-homolog 5 (imported) - garden snapdragon (fragment)

C:Species: Antirrhinum majus (garden snapdragon)

C:Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 20-Oct-2000

C:Accession: T52297

R:Cardon, G.H.; Hohmann, S.; Klein, J.; Nettesheim, K.; Saedler, H.; Huijser, P.

Gene 237, 91-104, 1999

A:Title: Molecular characterisation of the Arabidopsis SBP-box genes.

A:Reference number: Z25236; MUID:99453765; PMID:10524240

A:Accession: T52297

A:Status: Preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-419 <CAR>
A:Cross-references: EMBL:AJ011626; PIDN:CAB56578.1
A:Experimental source: cultivar Columbia; inflorescence
C:Genetics:
A:Gene: sp12
A:Map position: 5
C:Keywords: DNA binding; transcription factor

Query Match 19.2%; Score 392; DB 2; Length 419;
Best Local Similarity 31.5%; Pred. No. 4.2e-22;
Matches 116; Conservative 49; Mismatches 111; Indels 92; Gaps 16;

Qy 13 PQQAASG---GSSSTSSPSFG---GLMPGQKTYFED---GGGSG---SSSSG 53
Db PGEFAGIDTSPSLSPSSGDPVLGKKGKRTYEDFEVENAGLGIPVTLASSVSP 153
Qy 54 GSNRRVRGGSGGSGQIPRCVGECCMDLTNAKGYSHRVGVHAKTPVTVAGIEOR 113
Db 154 VKSKSI-----PQLQTPHCQVEGENDLSAKDYHRKRICENHSKPKVVGVERR 208
Qy 114 FGOCCSRFHQLPEPDLKSCRRRLAGHNERRRKQPA-----SLSTLARYGRIAPS 166
Db 209 FGOCCSRFHQLPEPDLKSCRRRLSDHARRKRPRTYDGKQVDFVNNRFAIHR 268
Qy 167 LYENDGAGNMGSGFLNGOIGMPPSRITLTRY--MRPVS--PSMOINPMNVFSGSGVG 223
Db 269 -----SBEKTIWPSKXVPERVLMPOPAKEISDTEHNRGLDPTKTRAR 314
Qy 224 RTSFSSPEIMDTKLESYKIG--DSNCALSLSPHOPHDNNNNNNNNNNNNNNNNTRASSG 282
Db 315 AELFKEKEV---TISHWASQDLGALSLSL-----NSTTW-VSSS 352
Qy 283 FCPMTVTVMAQPPA---PSOHQYLNPPWFKNDNDMSVNLGRTYEDNQISSGTAM 339
Db 353 DQPRFTLDHHSNMQPVAHR-----SAAQLNSVSGVMDPPAVGPTAL 399
Qy 340 -----GEF 342
Db 400 HRNGVQGF 407

RESULT 8

152592
squamosa-promoter binding protein 6 [imported] - Arabidopsis thaliana
N:Alternate names: transcription factor sp16
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 24-Oct-2000 #sequence_revision 24-Oct-2000 #text_change 24-Oct-2000
C:Accession: T52592
R:Cardon, G.H.; Hoehmann, S.; Klein, J.; Nettesheim, K.; Saedler, H.; Huijser, P.
Gene 237, 91-104, 1999
A:Title: Molecular characterisation of the Arabidopsis SBP-box genes.
A:Reference number: Z25236; MUID:99453765; PMID:10524240
A:Accession: T52592
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-405 <CAR>
A:Cross-references: EMBL:AJ011643; PIDN:CAB56595.1
A:Experimental source: cultivar Columbia
C:Genetics:
A:Gene: sp16
A:Map position: 2
C:Keywords: DNA binding; transcription factor

Query Match 18.4%; Score 374.5; DB 2; Length 405;
Best Local Similarity 32.4%; Pred. No. 8.4e-21;
Matches 122; Conservative 38; Mismatches 139; Indels 77; Gaps 14;

Qy 15 QAESGSSSTSSPSFGGLMFQKTYFE-----DGGGSGSSSSSGSRNRRVAGGSGGSG 69
Db 62 BEDEENISSSKSFSSQEL--NRIDFKLRSLDLGNDDDTSSRGAPALPSKKSRAASNLCS 119

Qy 70 QIPRCVGECCMDLTNAKGYSHRVGVHAKTPVTVAGIEORFQQCCSRFHQLPEFDL 129
Db 120 QNPCLQYVGCSDKSLSKDYKRRHVRCEBASKTSVTVNGLEQGFQQCCSRFHFLSSFFD 179
Qy 130 EKRCRRRLAGHNERRRKPPQASLSVLAARYGRVLAFLYENGDMGNGSGFLNGEIQMPS 189
Db 180 GKRCRRRLAGHNERRRK--PAFYFLPKRR-----KLRTSQDVANKFLSNSLVLP 232
Qy 190 S-----RTLDTRYMR--RPVS--SPSMOINPMNVFSGSGVGGRTSFPSPIMDTKL 237
Db 233 SFPSSLYRVIVDEDDHRTSLRVSFKDEPTCMFPTN--EQNS--SRITYSKPAIYSTEV 287
Qy 238 ESYKIGD---SNCAISLISNPHQPHDNNNN-----NNNNNNNNNNTRASSGSGF 284
Db 288 SSIMDLHETASRSTRALSLISAQSOQLSKFPPTTESITQPNQNLHSSST----- 339
Qy 285 PMYTVMAQPPAPSOHQYLNPPWFKNDNDMSVNLGRTYEDNQISSGTAMGEFL 344
Db 340 -----DYHQWEOPLMTIDPKTNSAG-----SSCKGKGTSTVDLLQL 376
Qy 345 SDH--HHOSRRQYMED 358
Db 377 SSHLQRIEQCRNRYGCD 392

RESULT 9

152596
squamosa promoter binding protein homolog 10 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 24-Oct-2000 #sequence_revision 24-Oct-2000 #text_change 24-Oct-2000
C:Accession: T52596
R:Cardon, G.H.; Hoehmann, S.; Klein, J.; Nettesheim, K.; Saedler, H.; Huijser, P.
Gene 237, 91-104, 1999
A:Title: Molecular characterisation of the Arabidopsis SBP-box genes.
A:Reference number: Z25236; MUID:99453765; PMID:10524240
A:Accession: T52596
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-396 <CAR>
A:Cross-references: EMBL:AJ011637; PIDN:CAB56589.1
A:Experimental source: cultivar Columbia; inflorescence
C:Genetics:
A:Gene: sp110
C:Keywords: DNA binding; transcription factor

Query Match 17.5%; Score 356; DB 2; Length 396;
Best Local Similarity 29.4%; Pred. No. 2e-19;
Matches 103; Conservative 40; Mismatches 97; Indels 110; Gaps 11;

Qy 3 MGSNSGPHG-----PGQA-ESGSGSTSSPSFGGLMFQKTYFEDGGGSGSSS 51
Db 90 LASQSPDSSNIDPLQVVKSTALEVPVIAESDLC--LKGKRTYSEBFMGRNNNDL 146
Qy 52 SGGSN-----RRVGGSGGSGQIPRCVGECCMDLTNAKGYSHRVGVHAKTPK 104
Db 147 SAVEMNLTPSVVARKTKKSCGQGMQVPRQCIDCELDLSSKDYHKKRVCETHSKCPK 206
Qy 105 VTVAGIEORFQCCSRFHQLPEPDLKSCRRRLAGHNERRRKPPQASLSVLAARYGRIA 164
Db 207 VVVGLERRFQCCSRFHAVSEFEKRSCKRKLSHNARRKQ---GVFLNSERVF 262
Qy 165 PSLYENDGAGNMGSGFLNGOIGMPPSRITLTRYMRPVS--PSMOINPMNVF----- 216
Db 263 -----DRQHTSMWN-----GLSLNTRSEKXTWGTYYETKPTQMSGFTLSFOR 308
Qy 217 -----QGSVGGGRTSFSFPIMDTKLESYKIG-----D 245
Db 309 GNGSEDQLFTGSTLSFAFQTSGGFGSAGKSNIQLPD-----KGVGCSGGLHSHD 359
Qy 246 SNCALSLISNPHQPHDNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 295
Db 360 FYSALSLST-----TSDSGIKHTFVAPPP 386

562361
 Squamosa-promoter binding protein 2 - garden snapdragon
 C:Species: Antirrhinum majus (garden snapdragon)
 C:Date: 23-Jul-1996 #sequence_revision 13-Mar-1997 #text_change 21-Jul-2000
 C:Accession: S62361
 R:Klein, J.; Seidler, H.; Huijser, P.
 Mol. Gen. Genet. 250, 7-16, 1996
 A:Title: A new family of DNA binding proteins includes putative transcriptional regulators
 A:Reference number: S62360; MUID:96158840; PMID:8569690
 A:Accession: S62361
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-171 <KLE>
 A:Cross-references: EMBL:X92079; NID:G1183863; PIDN:CAA63061.1; PID:G1183864

Query Match 16.2%; Score 330; DB 2; Length 171;
 Best Local Similarity 63.7%; Pred. No. 6.9e-18;
 Matches 65; Conservative 5; Mismatches 28; Indels 4; Gaps 2;

Qy 46 GSGSSSGGRNRRVGGGSGGQIPRCQVCGMDLTNAAGYYSRRHRCVGHSTKPKV 105
 Db 61 GGG---SSGKSEKHTATAGCGVVAQ-PCLVENCGADLNCCKYQGRHRCVGHAKAPV 116
 Qy 106 TVAGIEQRFCCQCSRFHQLPEFDLEKRSRRRLAGHNERRK 147
 Db 117 SVEGIMORFCCQCSRFHDLSEFDQTKRSRRRLAGHNERRK 158

RESULT 14
 T52607
 Squamosa promoter binding protein 5 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 24-Oct-2000 #sequence_revision 24-Oct-2000 #text_change 24-Oct-2000
 C:Accession: T52607
 R:Cardon, G.H.; Hoehmann, S.; Klein, J.; Nettesheim, K.; Seidler, H.; Huijser, P.
 Gene 237, 91-104, 1999
 A:Title: Molecular characterisation of the Arabidopsis SBP-box genes.
 A:Reference number: Z25236; MUID:99453765; PMID:10524240
 A:Accession: T52607
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: mRNA
 A:Residues: 1-181 <CAR>
 A:Cross-references: EMBL:AJ011610; PIDN:CAB56572.1
 A:Experimental source: cultivar Landsberg erecta; Inflorescence
 C:Genetics:
 A:Gene: ap1s
 A:Map position: 3
 C:Keywords: DNA binding; transcription factor

Query Match 16.2%; Score 330; DB 2; Length 181;
 Best Local Similarity 43.4%; Pred. No. 7.3e-18;
 Matches 75; Conservative 22; Mismatches 42; Indels 34; Gaps 7;

Qy 45 GSGSSSSGGRNR-----RVGGGSGGSGQIPR--COVEGCGMDLTNAAGYYSRRHRC 96
 Db 29 GMDGEEEDGDDCKRKVMERYRG--PSTIDVPSRLCQVDRCVLTLEAKQYRRHRC 85
 Qy 97 GVHSTPKVTVAIGIQRFCQCSRFHQLPEFDLEKRSRRRLAGHNERRKQPASLSVL 156
 Db 86 EVHAADSAATVAVGVRFCQCSRFHQLPEFDLEKRSRRRLAGHNERRK-----IS 138
 Qy 157 ASRYGRAPSLYENGAGMNGSFLGNOETIGWPSRTLDTRVMR--PVSSPSMQ 208
 Db 139 GDSFGE-----GSGRRGFSGLIQTE-----RNRVDRKLPMTNSSPK 176

RESULT 15
 H96793
 unknown protein FlAG6.18 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: H96793
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 ansen, N.F.; Hughes, B.; Huitzer, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;
 C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon, I.
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: H96793
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1020 <STO>
 A:Cross-references: GB:AE005173; NID:96642682; PIDN:AAF20262.1; GSPDB:GN00141
 C:Genetics:
 A:Gene: FlAG6.18
 A:Map position: 1

Query Match 16.1%; Score 328.5; DB 2; Length 1020;
 Best Local Similarity 29.7%; Pred. No. 7.4e-17;
 Matches 105; Conservative 36; Mismatches 104; Indels 109; Gaps 13;

Qy 55 RSNRRV-----GGSGGSGQIPRCQVCGMDLTNAAGYYSRRHRCVGHSTKPKVTAGI 110
 Db 91 RPSKKVRSGSPGSGGGGNYPKQVNDCKEDLSIADYHRRHRCVGHSTKPKVTALVGKQ 150
 Qy 111 EQRFCQCSRFHQLPEFDLEKRSRRRLAGHNERRKQP----- 150
 Db 151 MORFCQCSRFHLSFDEGRKSCRRLDGHNRRTKQPPAITSQVVALENRNTSNT 210
 Qy 151 -----ASISVLASRYGRAPSLYENGAGMNGSFLGNOETIGWPSRTLDTRVMRPPVSSP 205
 Db 211 NMDVALLTALVCAQGR-----NEATTNGS-----PQVPOREQ-LQILNKIKALP 255
 Qy 206 SWQINPMNVFQSGVGGRTSFSSPEIMDTLSEYKIGDSNCALSLSNPHOPHDNNN 265
 Db 256 L-----PNNLTS-----KL-----NNIGILARKNPEQSPNMPQ 284
 Qy 266 NNNNNNNNTW-----RASSGFGMTVTMAQPPAPSOHOYLPVWFVKDNDMSPVLN 321
 Db 285 NSMNGASSPTMDLALSLASLG-----SSAPEIAFLSGGFGNGKESNRT----- 331
 Qy 332 LGRYTEPDNCOISSGTAM-----GEFELSDHHQSRROYEDENTRA 363
 Db 332 --KLTSDDH--SATTSLSEKTLFEPFSGGERTSTNHSPOYSDSRQDTRS 380

Search completed: April 10, 2003, 13:54:35
 Job time : 23 secs

GenCore version 5.1.4 p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 10, 2003, 13:49:14 ; Search time 37 Seconds
(without alignments)
1350.513 Million cell updates/sec

Title: US-09-839-185-2

Perfect score: 2040
Sequence: 1 MEMGNSGPGHGPQAESEGG.....MEDENTRAYDSSHTNWSL 375Scoring table: BLOSUM62
Gapop 10.0 , Gapept 0.5

Searched: 908470 seqs, 13350620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 200000000Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2040	100.0	375	21	AA19952
2	2033	99.7	375	21	AA19952
3	2033	99.2	375	21	AA19952
4	2010.5	98.6	378	21	AA19952
5	2000.5	98.1	376	21	AA19952
6	1866	91.5	343	21	AA19952
7	1843.5	90.4	346	21	AA19952
8	863.5	42.3	350	21	AA19952
9	863.5	42.3	354	21	AA19952
10	693	34.0	289	21	AA19952

11	514.5	25.2	230	21	AA19952	Eucalyptus grandis
12	401	19.7	393	21	AA19952	Arabidopsis thaliana
13	401	19.7	393	21	AA19952	Arabidopsis thaliana
14	401	19.7	393	21	AA19952	Arabidopsis thaliana
15	401	19.7	425	21	AA19952	Arabidopsis thaliana
16	401	19.7	425	21	AA19952	Arabidopsis thaliana
17	401	19.7	425	21	AA19952	Arabidopsis thaliana
18	401	19.7	440	21	AA19952	Arabidopsis thaliana
19	396	19.4	359	21	AA19952	Arabidopsis thaliana
20	396	19.4	359	21	AA19952	Arabidopsis thaliana
21	394	19.3	336	21	AA19952	Arabidopsis thaliana
22	394	19.3	336	21	AA19952	Arabidopsis thaliana
23	392	19.2	387	21	AA19952	Arabidopsis thaliana
24	392	19.2	387	21	AA19952	Arabidopsis thaliana
25	392	19.2	387	21	AA19952	Arabidopsis thaliana
26	392	19.2	419	21	AA19952	Arabidopsis thaliana
27	392	19.2	419	21	AA19952	Arabidopsis thaliana
28	392	19.2	419	21	AA19952	Arabidopsis thaliana
29	375	18.4	257	21	AA19952	Arabidopsis thaliana
30	374.5	18.4	405	23	AA19952	Arabidopsis thaliana
31	354.5	17.4	376	21	AA19952	Arabidopsis thaliana
32	354.5	17.4	376	21	AA19952	Arabidopsis thaliana
33	354.5	17.4	376	21	AA19952	Arabidopsis thaliana
34	354.5	17.4	376	21	AA19952	Arabidopsis thaliana
35	354.5	17.4	389	21	AA19952	Arabidopsis thaliana
36	354.5	17.4	389	21	AA19952	Arabidopsis thaliana
37	354.5	17.4	389	21	AA19952	Arabidopsis thaliana
38	354.5	17.4	389	21	AA19952	Arabidopsis thaliana
39	354.5	17.4	389	21	AA19952	Arabidopsis thaliana
40	354.5	17.4	393	21	AA19952	Arabidopsis thaliana
41	354.5	17.4	393	21	AA19952	Arabidopsis thaliana
42	354.5	17.4	393	21	AA19952	Arabidopsis thaliana
43	351.5	17.2	349	21	AA19952	Arabidopsis thaliana
44	339.5	16.6	262	21	AA19952	Arabidopsis thaliana
45	339.5	16.6	262	21	AA19952	Arabidopsis thaliana

ALIGNMENTS

RESULT 1	
AA19952	AA19952 standard; Protein; 375 AA.
XX	
XX	AA19952;
XX	
XX	09-AUG-2000 (first entry)
XX	
XX	Arabidopsis SERK interacting protein #1 from clone 3A35.
XX	
XX	Somatic Embryogenesis Receptor Kinase; SERK; vegetative reproduction;
XX	apomixis; signal transduction; apomictic seed; somatic embryogenesis;
XX	Squamosa-promoter Binding Protein; SBP; transcription factor;
XX	transgenic plant.
XX	
XX	Arabidopsis thaliana.
XX	
XX	
XX	Key
XX	Region
XX	Location/Qualifiers
XX	1..58
XX	/note= "No corresponding codons in the nucleotide
XX	sequence"
XX	242..375
XX	/note= "No corresponding codons in the nucleotide
XX	sequence"
XX	
XX	W0200024914-A2.
XX	
XX	04-MAY-2000.
XX	
XX	20-OCT-1999;
XX	99NO-EP07972.
XX	
XX	22-OCT-1998;
XX	98GB-0023098.

PA (NOVS) NOVARTIS AG.
PA (NOVS) NOVARTIS-ERRINDUNGEN VERN GES MBH.
XX
XX Schmidt EDL, De Vries SC, Hecht VFG;
XX WPI; 2000-350753/30.
XX N-PSDB; AAD00300.
XX
PT Increasing vegetative production of a new plant generation through
PT apomixis, comprising transgenically expressing a gene encoding a
PT protein acting in the signal transduction cascade triggered by the
PT Somatic Embryogenesis Receptor Kinase
XX
XX Claim 4; Page 26-27; 45P; English.
XX
CC The patent discloses a method for increasing vegetative reproduction of
CC plants through seeds called apomixis, by transgenic expression of a gene
CC encoding a protein acting in the signal transduction cascade, triggered
CC by Somatic Embryogenesis Receptor Kinase (SERK). The protein physically
CC interacts with SERK, that is involved in the formation of extraneous
CC embryos from sporophytic cells, which can result in apomictic seeds.
CC This method is useful for increasing vegetative reproduction in vivo
CC through seeds or in vitro by somatic embryogenesis.
CC The present amino acid sequence is the SERK-interacting protein #1,
CC isolated from clone 3A35. It is obtained by RT-PCR amplification of
CC early stage Arabidopsis salique RNA. It is a member of the family of
CC Squamosa-promoter Binding Protein (SBP) transcription factors. SERK
CC transphosphorylates cytoplasmic SBP-transcription factors, after binding
CC of ligand, followed by nuclear translocation of the factors and binding
CC to specific regulatory DNA target sites on the genome.
XX
XX Sequence 375 AA;
SQ
Query Match 100.0%; Score 2040; DB 21; Length 375;
Best Local Similarity 100.0%; Pred. NO. 3.9e-176; Indels 0; Gaps 0;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MEMGNSGPHGHPQAEAGSGSTSSSFSQGLMFGQKTYFEDGGGSGSSSGGRNRRV 60
DB 1 MEMGNSGPHGHPQAEAGSGSTSSSFSQGLMFGQKTYFEDGGGSGSSSGGRNRRV 60
QY 61 RGGSGSGGQIIPRCVGGGMDLTNAKGYSHRVGVHSTKPYTVAGIEORFCQCSR 120
DB 61 RGGSGSGGQIIPRCVGGGMDLTNAKGYSHRVGVHSTKPYTVAGIEORFCQCSR 120
QY 121 FHQLEPFLERKSCRRRLAGHERRRRPPASLSTVLAHYGRVIASTVENGDAGNNGSFL 180
DB 121 FHQLEPFLERKSCRRRLAGHERRRRPPASLSTVLAHYGRVIASTVENGDAGNNGSFL 180
QY 181 GNOEIGWSSRLTDRVMRRPVSSPSWQINPMNVFSQGVGGRTSFSSEPEIMDKLEBY 240
DB 181 GNOEIGWSSRLTDRVMRRPVSSPSWQINPMNVFSQGVGGRTSFSSEPEIMDKLEBY 240
QY 241 KGIGDSCALSTLSNPHOPHDNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 300
DB 241 KGIGDSCALSTLSNPHOPHDNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 300
QY 301 OYLNPPWFKDNDNMSPTLNGRYTEPDNCOISGTAMGEFELSDHHOSRQYMEDN 360
DB 301 OYLNPPWFKDNDNMSPTLNGRYTEPDNCOISGTAMGEFELSDHHOSRQYMEDN 360
QY 361 TRAYDSSSHHTWMSL 375
DB 361 TRAYDSSSHHTWMSL 375
RESULT 2
AAG31717
ID AAG31717 standard; Protein; 375 AA.
XX AAG31717;
XX
DT 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 38136.
DE Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.
XX Arabidopsis thaliana.
OS
XX EP1033405-A2.
PN
XX
PD 06-SEP-2000.
XX
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
KM 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
XX 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
XX 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
XX 07-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
XX 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136352.
PR 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
XX 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
XX 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139454.
XX 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
XX 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
XX 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
XX 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
XX 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
XX 18-JUN-1999; 99US-0139763.


```
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140635.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148177.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149172.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.

PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160747.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 99.7%; Score 2033; DB 21; Length 375;
Best Local Similarity 99.7%; Pred. No. 1.7e-175;
Matches 374; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MEMSNSGPHGPGQASGSGSSTSSFSGGLMPGQKIYFEDGGGSGSSGGRSNRV 60
Db 1 MEMSGSGPHGPGQASGSGSSTSSFSGGLMPGQKIYFEDGGGSGSSGGRSNRV 60
QY 61 RGGSGSGGQIPRCQVBCGMDLTNAKYYSRHRVCGVSHKTPYTAGIBORFCCQCSR 120
Db 61 RGGSGSGGQIPRCQVBCGMDLTNAKYYSRHRVCGVSHKTPYTAGIBORFCCQCSR 120
QY 121 FHOLPEPDLERKSCRRLAGNERRRKPQPASLSVLASRYRIAPSLYENDAGMNGSFL 180
Db 121 FHOLPEPDLERKSCRRLAGNERRRKPQPASLSVLASRYRIAPSLYENDAGMNGSFL 180
QY 181 GNOEIGMPSRTLDTRVWRPVSPPWQINPMVFSQSGYGGRTSFSPEIMDTKLESY 240
Db 181 GNOEIGMPSRTLDTRVWRPVSPPWQINPMVFSQSGYGGRTSFSPEIMDTKLESY 240
QY 241 KGIGDSNCALSLSNPHQPHDNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 300
Db 241 KGIGDSNCALSLSNPHQPHDNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 300
```

```
Qy 301 QYLNPPWVFKDNDNMDSPVLTNIGRYTEPDNCQISSGTAMGEFELSDHHQSRQYMEDEN 360
Db 301 QYLNPPWVFKDNDNMDSPVLTNIGRYTEPDNCQISSGTAMGEFELSDHHQSRQYMEDEN 360
Qy 361 TRAYSSSSHHHTWWSL 375
Db 361 TRAYSSSSHHHTWWSL 375

RESULT 3
AAG31718
ID AAG31718 standard; Protein; 373 AA.
XX
AC AAG31718;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 38137.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 01-APR-1999; 99US-0126785.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137503.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.

PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 26-JUN-1999; 99US-0140695.
PR 29-JUN-1999; 99US-0140823.
PR 30-JUN-1999; 99US-0140981.
PR 01-JUL-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 19-JUL-1999; 99US-0144335.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 28-JUL-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147392.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
```

PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	14-AUG-1999;	99US-0148684.
PR	15-AUG-1999;	99US-0149366.
PR	16-AUG-1999;	99US-0149175.
PR	17-AUG-1999;	99US-0149426.
PR	18-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	26-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155133.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	18-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160880.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	23-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.

Query Match 99.2%; Score 2023; DB 21; Length 373;
 Best Local Similarity 99.7%; Pred. No. 1,3e-17;
 Matches 372; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db	1	MGSNSGPGHGCGQABSGSSSTRESSFSGCLMFQGXITYFEDGGGGSGSSSGGRSNRRVKG	60
Qy	63	GSGSGSGGIPFCQYEGGCGMDLTNAKGYYSRRVCGVASKTPKTVAGIEORFCQCSRFH	122
Db	61	GSGSGSGGIPRCQYEGGCGMDLTNAKGYYSRRVCGVASKTPKTVAGIEORFCQCSRFH	120
Qy	123	QLPEFDLEKRSRRRLAGHNERRRKPPASLSVLASYGRIAPSLVYENGDAQMGNSFLGN	182
Db	121	QLPEFDLEKRSRRRLAGHNERRRKPPASLSVLASYGRIAPSLVYENGDAQMGNSFLGN	180
Qy	183	QEIGMPSSRTLDTVMMRPVSPSPMOINPMNVFSGSVGGGRSPSSPELMDTKLESYKG	242
Db	181	QEIGMPSSRTLDTVMMRPVSPSPMOINPMNVFSGSVGGGRSPSSPELMDTKLESYKG	240
Qy	243	IGDSNCALSTLSNPHOPDHNNNNNNNNNNNTWRASGFGPVTYMAOPPPAPSOHQY	302
Db	241	IGDSNCALSTLSNPHOPDHNNNNNNNNNNNTWRASGFGPVTYMAOPPPAPSOHQY	300
Qy	303	LNPPWFVKDNDNDSPVLTNGRYTEPNCCQISSGTAMGEFELSDHHHQSRQYMEDENTR	362
Db	301	LNPPWFVKDNDNDSPVLTNGRYTEPNCCQISSGTAMGEFELSDHHHQSRQYMEDENTR	360
Qy	363	AYDSSSHHTNWSL 375	
Db	361	AYDSSSHHTNWSL 373	
RESULT 4			
ID	AA616706	AA616706 standard; Protein; 378 AA.	
XX	AC	AA616706;	
XX	DT	17-OCT-2000 (first entry)	
DE	XX	Arabidopsis thaliana protein fragment SEQ ID NO: 17455.	
KW	KM	Protein identification; signal transduction pathway; metabolic pathway;	
KW	KM	hybridisation assay; genetic mapping; gene expression control; promoter;	
OS	XX	Arabidopsis thaliana.	
PN	XX	EP1033405-A2.	
PD	XX	06-SEP-2000.	
PF	XX	25-FEB-2000; 2000EP-0301439.	
PR	XX		
PR	05-FEB-1999;	99US-0121825.	1
PR	05-MAR-1999;	99US-0123180.	
PR	09-MAR-1999;	99US-0123548.	
PR	23-MAR-1999;	99US-0125788.	
PR	25-MAR-1999;	99US-0126264.	
PR	29-MAR-1999;	99US-0126785.	
PR	01-APR-1999;	99US-0127462.	
PR	06-APR-1999;	99US-0128234.	
PR	08-APR-1999;	99US-0128714.	
PR	16-APR-1999;	99US-0129845.	
PR	19-APR-1999;	99US-0130077.	
PR	21-APR-1999;	99US-0130449.	
PR	23-APR-1999;	99US-0130510.	
PR	23-APR-1999;	99US-0130891.	
PR	28-APR-1999;	99US-0131449.	
PR	30-APR-1999;	99US-0132048.	
PR	30-APR-1999;	99US-0132407.	
PR	04-MAY-1999;	99US-0132484.	
PR	05-MAY-1999;	99US-0132485.	
PR	06-MAY-1999;	99US-0132486.	
PR	06-MAY-1999;	99US-0132487.	
PR	07-MAY-1999;	99US-0132863.	
PR	11-MAY-1999;	99US-0134256.	
PR	14-MAY-1999;	99US-0134218.	

PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0136021.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140921.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.

PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147260.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156456.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.

PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 98.6%; Score 2010.5; DB 21; Length 378;
Best Local Similarity 98.7%; Pred. No. 1,8e-173;
Matches 373; Conservative 0; Mismatches 2; Indels 3; Gaps 1;

QY 1 MEMSNGSGPGHGPQAESEGSSTSSSSFGGLMFGQKTYFEDGGGSSSSGGRSNRV 60
DB 1 MEMSNGSGPGHGPQAESEGSSTSSSSFGGLMFGQKTYFEDGGGSSSSGGRSNRV 60
QY 61 RGGSGSGSGQIPRCQVEGCGMDLTNAKGYSRHRVCGVHSKTPVTYAGIBQRCQCCSR 120
DB 61 RGGSGSGSGQIPRCQVEGCGMDLTNAKGYSRHRVCGVHSKTPVTYAGIBQRCQCCSR 120
QY 121 FHOUPEDLEKSCRRRLAGNERRRKQPASLSYLASRYGRIPASLYENDAGNNGSFL 180
DB 121 FHOUPEDLEKSCRRRLAGNERRRKQPASLSYLASRYGRIPASLYENDAGNNGSFL 180
QY 181 GNOEIGMSSSTLTTRVWRVSSPQINPMNVFSQSGVGGRTSFSPEIMTKLESY 240
DB 181 GNOEIGMSSSTLTTRVWRVSSPQINPMNVFSQSGVGGRTSFSPEIMTKLESY 240
QY 241 KGIDSNCAISLISNPHQPD---NNNNNNNNNNNTWRASGFGPMVTMAQPPAP 297
DB 241 KGIDSNCAISLISNPHQPD---NNNNNNNNNNNTWRASGFGPMVTMAQPPAP 297
QY 298 SOHOYLNPWVFKONDNMSPVNLIGRYTEPDNCOISSGTAMGEFELSDHHOSRQYME 357
DB 301 SOHOYLNPWVFKONDNMSPVNLIGRYTEPDNCOISSGTAMGEFELSDHHOSRQYME 360
QY 358^{*}DENTRAYDSSSHHTNWSL 375
DB 361 DENTRAYDSSSHHTNWSL 378

RESULT 5

AAAG16707 standard; Protein: 376 AA.

AAAG16707;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 17456.

KW Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 23-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.

```
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144332.
PR 20-JUL-1999; 99US-0144332.
PR 20-JUL-1999; 99US-0144332.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145226.
PR 27-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147309.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.

PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161982.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 98.1%; Score 2000.5; DB 21; Length 376;
Best Local Similarity 98.7%; Pred. No. 1.5e-172;
Matches 371; Conservative 0; Mismatches 2; Indels 3; Gaps 1;

QY 3 MGSNSGPHGPGQAEBSGSSSTESSFSQGLMFQKIFYEDGGSGSSSGGRNRRVG 62
DB 1 MGSNSGPHGPGQAEBSGSSSTESSFSQGLMFQKIFYEDGGSGSSSGGRNRRVG 60
QY 63 GSGSGSGQIPRCQYBEGCMTLTNAKGYRRVGVHSHKTKYTVAGIEORFCQCCSFH 122
DB 61 GSGSGSGQIPRCQYBEGCMTLTNAKGYRRVGVHSHKTKYTVAGIEORFCQCCSFH 120
QY 123 QLPFEDLEKRSRRRLAGHNERRRKPOPASLSTVLSRYGRITAPSLYENGDAKMGSLFGN 182
DB 121 QLPFEDLEKRSRRRLAGHNERRRKPOPASLSTVLSRYGRITAPSLYENGDAKMGSLFGN 180
QY 183 QEIGWPSRTLTDTVRMRPVSSPSMQINPMNVFSQSVGGRTSFSSEIMDTKLESYKG 242
DB 181 QEIGWPSRTLTDTVRMRPVSSPSMQINPMNVFSQSVGGRTSFSSEIMDTKLESYKG 240
QY 243 IGDNSCALSLSNPHQPHD--NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 299
DB 241 IGDNSCALSLSNPHQPHDNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 300
QY 300 HOYLNPWFVKONDNDMSPVNLGRYTEPDNCOISSGTAMGEFELSDHHOSRQYMEDE 359
DB 301 HOYLNPWFVKONDNDMSPVNLGRYTEPDNCOISSGTAMGEFELSDHHOSRQYMEDE 360
QY 360 NTRAYDSSSHHTNWSL 375
DB 361 NTRAYDSSSHHTNWSL 376

RESULT 6
AAG31719
ID AAG31719 standard; Protein; 343 AA.
AC AAG31719;
XX
XX 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 38138.
XX
XX Protein identification: signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
```

KM termination sequence.
XX Arabidopsis thaliana.
XX
XX EPI033405-A2.
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135533.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.

PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145213.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149920.
PR 25-AUG-1999; 99US-0150586.
PR 25-AUG-1999; 99US-0150586.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.

PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 91.5%; Score 1866; DB 21; Length 343;
Best Local Similarity 99.7%; Pred. No. 1.9e-160;
Matches 342; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 33 MFGQKTYFEDGGGSGSSSGGSRNRRVVGSGSGSQI PRQVBEQCGMDLTNAKYYSR 92
DB 1 MFGQKTYFEDGGGSGSSSGGSRNRRVVGSGSGSQI PRQVBEQCGMDLTNAKYYSR 60
QY 93 HRVCGVHAKTPKVTYVAGI BQRFCCQCCSRFHQI PEEDLEKRSRRRLAGHNERRRKQPAS 152
DB 61 HRVCGVHAKTPKVTYVAGI BQRFCCQCCSRFHQI PEEDLEKRSRRRLAGHNERRRKQPAS 120
QY 153 LSVLASRYGRITAPSLYXMGDAGMNSFLGNORIGWPSRTLDLTRVMRRPVSSPSQIINPM 212
DB 121 LSVLASRYGRITAPSLYXMGDAGMNSFLGNORIGWPSRTLDLTRVMRRPVSSPSQIINPM 180
QY 213 NVFSQGSYVGGRTRTSSPEIMDTKLESYKIGIDSNCAISLISNPHQPHDNNNNNNNNNN 272
DB 181 NVFSQGSYVGGRTRTSSPEIMDTKLESYKIGIDSNCAISLISNPHQPHDNNNNNNNNNN 240
QY 273 NNNTWRASSGFGPMTVTMAQPPAPASQHOYLNPPVFKDNDNDMPVLNLGTYTEPDNCQ 332
DB 241 NNNTWRASSGFGPMTVTMAQPPAPASQHOYLNPPVFKDNDNDMPVLNLGTYTEPDNCQ 300
QY 333 ISSGTAMGEFELSDHHHOSRQYMEDENTRAYDSSSHHTNWSL 375
DB 301 ISSGTAMGEFELSDHHHOSRQYMEDENTRAYDSSSHHTNWSL 343

RESULT 7
AAG16708
ID AAG16708 standard; Protein; 346 AA.
XX
AC AAG16708;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 17457.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
PD
XX 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 29-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137528.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.


```

PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139819.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147182.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149358.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.

```

```

PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

Query Match 90.4%; Score 1843.5; DB 21; Length 346;
 Best Local Similarity 98.6%; Pred. No. 2.1e-158;
 Matches 341; Conservative 0; Mismatches 2; Indels 3; Gaps 1;

```

QY 33 MFGOKTYFEDGGGSSGSSGGRSNRRVRRGGSGSGSQIPRCQVEGGMGLTNAKGYYSR 92
Db 1 MFGOKTYFEDGGGSSGSSGGRSNRRVRRGGSGSGSQIPRCQVEGGMGLTNAKGYYSR 60
QY 93 HRVGVHAKTKPTVAGIEORFCQOCRSRPHQLEPFDLKRSRRRLAGHNERRRKQPAS 152
Db 61 HRVGVHAKTKPTVAGIEORFCQOCRSRPHQLEPFDLKRSRRRLAGHNERRRKQPAS 120
QY 153 LSVLASRYGRIAPSLYENGGAGWNGSFLGNQETIGWPSRRLDLTRWARRPVSSSQINPM 212
Db 121 LSVLASRYGRIAPSLYENGGAGWNGSFLGNQETIGWPSRRLDLTRWARRPVSSSQINPM 180

```

Qy 213 NVFSGSVGGRTSFSSBEINDTKLESYKGI GDSNCALSLSPHQPHD---NNNNNNN 269
Db 181 NVFSGSVGGGCTSFSSBEINDTKLESYKGI GDSNCALSLSPHQPHDNNNNNNNN 240
Qy 270 NNNNNNTVRASSGCGPMVTVMQAPPAPSOHQYLNPPVFKDNDNDMSPVNLGRYTEPD 329
Db 241 NNNNNNTVRASSGCGPMVTVMQAPPAPSOHQYLNPPVFKDNDNDMSPVNLGRYTEPD 300
Qy 330 NCOISSGTAMGEFELSDHHOSRQRYMEDENTRAYDSSSHHTNMSL 375
Db 301 NCOISSGTAMGEFELSDHHOSRQRYMEDENTRAYDSSSHHTNMSL 346
RESULT 8
AAG22917
ID AAG22917 standard; Protein: 350 AA.
AC AAG22917;
XX 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 26028.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX
XX EPI033405-A2.
XX
XX 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 28-APR-1999; 99US-0130891.
XX 30-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 04-MAY-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 06-MAY-1999; 99US-0132487.
XX 07-MAY-1999; 99US-0132863.
XX 11-MAY-1999; 99US-0134255.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134768.
XX 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139889.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 22-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 26-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.

PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0135124.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135253.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139889.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.

PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145813.
PR 27-JUL-1999; 99US-0145818.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147202.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 20-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152263.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.

PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 42.3%; Score 863.5; DB 21; Length 354;
Best Local Similarity 52.7%; Pred. No. 1,1e-69;

Matches 204; Conservative 32; Mismatches 86; Indels 65; Gaps 17;

QY 14 GQAESGQ-SSTESSFSGGLMFGQKIVFEDQGGSGSSSGGRNRVRGGSGSGQIP 72
DB 8 GQAESGSSSTESSLSGGLRFGQKIVFEDG-----SGSRKNRV--NTVAKSSTA 57
QY 73 RCQVEGCGMDLTNAKGYSSRRVCGVHSKTPKVTAGIEQPCQCSRFHQLPEFDLEKR 132
DB 58 RCQVEGCGMDLTNAKGYSSRRVCGVHSKTPKVTAGIEQPCQCSRFHQLPEFDLEKR 117
QY 133 SCRRRLAGHNERRRKPPQASLSVLSRYGRAPSLYENGAGMNGSFLGNGEIGWPSERT 192
DB 118 SCRRRLAGHNERRRKPPQASLSVLSRYGRAPSLYENGAGMNGSFLGNGEIGWPSERT 175
QY 193 LDTVRMRPVSSPFWQNP-----NMVFSQSGVGGRTSFS-S-PRIMDTKLESYKIG 244
DB 176 ----VMGRP---GWMQINPVRKETHPHNVLSHGS-----SFTTCPEMINNST----- 217
QY 245 DSNCAISLISNPHOPHDNNNNNNNNNNNTWPASSGFGFM-----TWTMAQP----- 293
DB 218 DSSCALSLISNSYPIH-----QQQLQPTVTWTPSSGFSMSISFSDKVTMAQPPPISTH 271
QY 294 -PPAPSOHOVLPWPV-----KDDNDMSPYLANIGRTPEPNCQISSGTMGEFELSDH 348
DB 272 QPPISTHQVLSQVWEVIAGKSNHYSVPV---SQISEPADFOISNGTWTGFEELY-LH 327
QY 349 HQSRROYMEDENTRAYDSSSHHTWMSL 375
DB 328 QQVYLKQIMEPEENTRAYDSSPHFWMSL 354

RESULT 10

ID AAG22918 standard; Protein; 289 AA.

AC AAG22918;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 26029.

KW Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.

OS Arabidopsis thaliana.

PN EPI033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 28-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.

XX basic helix-loop-helix zipper; homeotic; homeodomain; homeobox; MADS;
KM homeodomain zipper; LIM domain; AP2; ERBBs; zinc finger domain;
XX type 2 Cys2His2; CCAAT box element; MYB.
OS Eucalyptus grandis.
XX MO200053724-A2.
XX 14-SEP-2000.
PD 09-MAR-2000; 2000MO-US06112.
XX 11-MAR-1999; 99US-0266513.
XX 18-AUG-1999; 99US-0149485.
XX (GENE-) GENESIS RES & DEV CORP LTD.
PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
XX Wood M, McGrath A, Shenk MA, Glenn M,
XX WPI; 2000-579369/54.
XX
XX New isolated polynucleotide encoding a plant transcription factor for
PT producing a plant e.g. a woody plant, preferably eucalyptus or pine,
PT having modified gene expression or modified activity of a polypeptide
PT -
XX
XX Claim 8; Page 612; 747pp; English.
XX
XX The present invention relates to novel plant transcription factors from
CC Eucalyptus grandis or Pinus radiata. The present sequence is one such
CC transcription factor. The transcription factor may be used to produce a
CC plant having modified gene expression such as a woody plant e.g. a
CC eucalyptus, pine, acacia, poplar, sweetgum, teak, or mahogany species or
CC to modify the activity of a polypeptide in a plant. The transcription
CC factors of the present invention are members from the following families
CC of regulatory proteins: bZIP, bZIP family of G-box binding factors, basic
CC helix-loop-helix zipper, homeotic/homeodomain/homeobox/MADS, homeodomain
CC zipper, LIM domain, AP2 and ERBBs, zinc finger domains of type 2
CC Cys2His2, CCAAT box elements and MYB.
XX
XX Sequence 230 AA;
SQ
Query Match 25.2%; Score 514.5; DB 21; Length 230;
Best Local Similarity 57.1%; Pred. No. 2,3e-38;
Matches 112; Conservative 15; Mismatches 40; Indels 29; Gaps 6;
QY 10 GHGPGQAEAGSGSTSSSFGGLMFGQKTYFD-----GGGSGSGSSS--- 52
DB 4 GSGSWTTSSGSSSPPLSLINGLTKRGQKTYFQNNNSNNMAAPKNGSGSGSSSAADPA 63
QY 53 --GGRSNRRV---GGSGG--QSGQIRCOYEGCGMLTNAGKYSHRRCGVSKTPK 104
DB 64 PGSGTPPKKVASAGGGCGAIOGQPPRCOVGCGRVLSAKAYSHRKVCGMHSKAT 123
QY 105 VTVAGIEORFCQCGSRFPQLPEFDLEKRSRRRLAGHERRRKKPOASTSVLASRYGRIA 164
DB 124 VIVAGIEORFCQCGSRFPQLTFPDGKRSRRRLAGHERRRKKPPGSL--LSRYGRLQ 181
QY 165 PSLYENGAGNNGSFL 180
DB 182 SSIFEN--TRVGSFL 195
RESULT 12
AAG12773
ID AAG12773 standard; Protein; 393 AA.
XX
XX AAG12773;
AC
XX 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 12015.

XX Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX BP1033405-A2.
XX 06-SEP-2000.
PD 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 28-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 30-APR-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 06-MAY-1999; 99US-0132487.
XX 07-MAY-1999; 99US-0132863.
XX 11-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134768.
XX 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138847.
XX 14-JUN-1999; 99US-0139119.
XX 16-JUN-1999; 99US-0139452.
XX 16-JUN-1999; 99US-0139453.
XX 17-JUN-1999; 99US-0139452.
XX 18-JUN-1999; 99US-0139454.
XX 18-JUN-1999; 99US-0139455.
XX 18-JUN-1999; 99US-0139456.
XX 18-JUN-1999; 99US-0139457.
XX 18-JUN-1999; 99US-0139458.
XX 18-JUN-1999; 99US-0139459.
XX 18-JUN-1999; 99US-0139460.
XX 18-JUN-1999; 99US-0139461.
XX 18-JUN-1999; 99US-0139462.
XX 18-JUN-1999; 99US-0139463.
XX 18-JUN-1999; 99US-0139763.
XX 18-JUN-1999; 99US-0139817.
XX 22-JUN-1999; 99US-0139899.

PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 06-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147499.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.

PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152263.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158332.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159337.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161362.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 19.7%; Score 401; DB 21; Length 393;
Best Local Similarity 31.3%; Pred. No. 9e-28;
Matches 121; Conservative 50; Mismatches 108; Indels 108; Gaps 17;

QY 13 PGQAEAG---GSSTESSFSG---GLMFGQKIYFED-----GGGSGS-----SSSGS 53
DB 65 PGEPAKGIDTSPTELEISFGSGDPVLGLKLGKRTYFEDFEVENAKGGLPVSLASSSVSP 124
QY 54 GRSNRVYGGSGSGSGQIPRCQVEGCGMDLTNAKGYSRHRVCVHSTPKYTVAGIEQR 113
DB 125 VKKSKSI-----PQRLQPHQVQEGCNLDLSSAKDYHRKHICENHSKFPKVVVSGVRR 179
QY 114 FCQCCSRPHQLPEDLEKRSRRRLAGNERRRKQPA-----SLSVASRYGR1APS 166
DB 180 FCQCCSRFHCLSEFDEKRSRRRLSDNARRKKNPGRITDGKRPQVDFVNRPAL1HPR 239
QY 167 LYENDAGMNSFLGNQELIGPSSRTLDTRYMRPVPSPMQINPMVYFSGSVGG---G 223
DB 240 SEE-----KFL-----WPSKFPVPSRVL---MPOPATETISNKLFTHSRFGLLDP 282
QY 224 RTSFSPFIMDTK---LESYKIG-G-DSNCALSLSLNPHQPHDNNNNNNNNNNNTWRA 279
DB 293 KTKSARALELFSKEKVTYISSHMGASODLDGALSLIS-----NSTTW-V 323
QY 280 SSGFGPMVTYMAQPPAPDSQHQYLANPWFVFKDNDNDSVULN-----LGRYTEPUNC 331


```
Db 324 SSSDPREFITDHP-----SSNLOPVANRGAQISSVGYWQDPP 365
Qy 332 QISSGTAMGEFELSDHHQSRQYMED 358
   366 AVEGPTAL-----HNGAGQNNEN 384

RESULT 13
AAG14392
ID AAG14392 standard; Protein; 393 AA.
XX AAG14392;
AC
XX 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 14236.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX
XX EPI033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126254.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 15-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 28-APR-1999; 99US-0130891.
XX 30-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 07-MAY-1999; 99US-0132487.
XX 11-MAY-1999; 99US-0132863.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 18-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134768.
XX 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138847.
XX 14-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 15-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
```

```
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162142.

Query Match 19.7% Score 401, DB 21, Length 393;
Best Local Similarity 31.3%; Pred. No. 9e-28;
Matches 121; Conservative 50; Mismatches 108; Indels 108; Gaps 17;
```

```
QY 13 PQQAESG---GSSSTSSSSG---GLMFGOKIYFPD-----GGGSG-----SSSSG 53
DB 65 PGEFAKIDTSPLELISFSGSDPVLGLKLGKRTYFEDFWEVENAKGLGLPVSILASSVSP 124
```

```
QY 54 GRSNRVRVGGSGSGQIPRCQVGGCMGLTNAGYVSRHRRVCGYHSTKTPKVTYAGIEOR 113
DB 125 VKKSKSI-----PQRLQTPQVGGCNLDLSAKDYRKHRIICENHSKFPKVVVSGVER 179
QY 114 FCCOCSRPHOLPEPDLDEKRCRRRLAGNERRRKPQPA-----SLSVASRYGRIAPS 166
DB 180 FCCOCSRPHCLSEFDEKRCRRRLSDHNAKRKPNRGRTYDGKPYQDFYWNRFALHPR 239
QY 167 LYENGDAKMGSGFLGNQOIEIMPSSRTLDTRVRRPVSSPQMGINPMNVFSGSVYG---G 223
DB 240 SEE-----KFL-----WPSKPVPSRYL---MPQPAKTEISNKLFTHSRFGLLDP 282
QY 224 RTSGFSSPEINDTK---LESTKIGC-DSCNCAISLISNPHQHDNNNNNNNNNNNTWRA 279
DB 283 KTKSARAELFSKEKYTITISHMGASQDLGALSTLS-----NSTTW-V 323
QY 280 SSGFPMVTVMAPPPAPSOHQYLNPPWVFKDNDNMSPYLN-----LGRYTEPDNC 331
DB 324 SSSQPRRFTLDHHP-----SSNIQPVANRGAQLSSVSGYWGQDPP 365
QY 332 QISSGTAMGEFELSDHHHOSRQYMED 358
DB 366 AVEGPTAL-----HRNGAGQFNEN 384

RESULT 14
AAG44414
ID AAG44414 standard; Protein; 393 AA.
AC AAG44414;
XX
XX 18-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 55632.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 23-APR-1999; 99US-0130891.
XX 28-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 30-APR-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 06-MAY-1999; 99US-0132487.
XX 07-MAY-1999; 99US-0132863.
XX 11-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
```

PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140654.
PR 24-JUN-1999; 99US-0140655.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144351.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144332.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.

PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.

PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161933.
PR 29-OCT-1999; 99US-0162142.

Query Match 19.7%; Score 401; DB 21; Length 393;
Best Local Similarity 31.3%; Pred. No. 9e-28;
Matches 121; Conservative 50; Mismatches 108; Indels 108; Gaps 17;

QY 13 PGQAEAG-SSSTSSSFSG-GLMFGQKIYFED-GGGGSG-SSSSG 53
D6 65 PGEPAKGDITPTLELSFGSDPVLGLTKRITYEDWEVENAKGLPVLSSSVSP 124
QY 54 GRSNRVRVGGSGSGGQIPRCQVSGGMDLTNAKYRHRVCCVHSTKPTVTAGIQR 113
D6 125 VKKSKSI-PQLQTPHCQVEGCNLDLSAKDYHKKRHCENHSKFPKVVSVGERR 179
QY 114 FCQCCSRPHOLPEPDLERKSGRRRLAGNERRRKQPA-SLSVLASRYGRIAPS 166
D6 180 FCQCCSRPHCLSEPEDEKRSRRRLSDHARRKPNPRTTDGKQVDFVMMRFLIHR 239
QY 167 LYENGDAAGNGSFLGNQRIQWPSRTLDTRYVRRPVSPPQINPMNVFSQSVGG-G 223
D6 240 SEF-KFL-MPSKRPVSRVL-MPPAKTEISNKLFTESRGLDDP 282
QY 224 RTSSSPPEIMDTK-LESYKIG-DSNCAISLSNPQHOPHNNNNNNNNNTTARA 279
D6 283 KTKSARALFSEKEKVTISSHMGASQDLDAISLIS-NSTTW-V 323
QY 280 SSGRPMTVTMAQPPAPASQOYVLPVWFKONDMSVPLN-LGRYTEPDNC 331
D6 324 SSSDQPRPRTLDHP-SNLDQVARRSAQSLSSVGYWQDPP 365
QY 332 QISSGTAGEFELSDHHQSRQYMED 358
D6 366 AVEGFTAL-HNGAQGFEN 384

RESULT 15
AAG12772
ID AAG12772 standard; Protein; 425 AA.
XX AAG12772;
AC
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 12014.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
FN EP1033405-A2.
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131149.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139859.
PR 23-JUN-1999; 99US-0140033.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142380.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.

PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145102.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 25-JUL-1999; 99US-0145276.
 PR 26-JUL-1999; 99US-0145291.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.

PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 19.7%; Score 401; DB 21; Length 425;
 Best Local Similarity 31.3%; Pred. No. 1e-27;
 Matches 121; Conservative 50; Mismatches 108; Indels 108; Gaps 17;

QY 13 PQQAESG---GSSTESSFSG---GIMFGQKTYFED---GGGSG---SSSG 53
 DB 97 PGEPAKIDTSPTELSFSGDPEVLGKLGKTYFEDFWEVENAKGIGLPSLASSVSP 156
 QY 54 GRSNRVRGGSGSGQIPPCQVEGCGMDLTNAKGYRRVCGVSKTKYTVAGIEOR 113
 DB 157 VKKSKST---PQRLQTPHCQVEGCNLDLSAKDYHRKRIICENHSKFEKVVSGVER 211
 QY 114 FCOQCSFHOLPEFDELEKRCRRRLAGHNERRRKPOPA-----SISVLASRYGRIAPS 166
 DB 212 FCOQCSFHCLSEFDEKRCRRRLSDHNAKRRKPNRGRTYDGPQVDYFWNRFALIHPR 271
 QY 167 LYENGDAWNGSFLGNQELGWPSSRTLDTRVMRRPVSSPSWQJNPMNVFSQSGVG---G 223
 DB 272 SEB-----KFL-----WPSKVPVSRVL---MPQAKTEISNLTPEHSRFGILD 314
 QY 224 RTSFSSPEIMDTK---LESYKIG-DSNCALSLISNPHQHDNNNNNNNNNNNTWRA 279
 DB 315 KTKSARAELEFSKEKVTITSHMGASODIDGALSLS-----NSTTW-V 355
 QY 280 SSGFGPMTVMAQPPAPSQHOYLNPWFVKDNDNDMSPVLN-----LGRYTEPDNC 331
 DB 356 SSSDQPRRFTLIDHNP-----SSNIQPVANSAAQLSVSGYMQDPDP 397
 QY 332 QISSGTAMGEFELSDHHDGRRQYMED 358
 DB 398 AVEGPAL-----HRNGAGQENEN 416

Search completed: April 10, 2003, 13:53:04
 Job time : 39 secs

Qy 315 DMSFVNLGRYTEPNCOISGTMGEFELSDHHQSRQYMEDETRAYDSSSHHTN 372
Db 462 NNNNNNNNNNNNNNNSSNTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 519

RESULT 2

US-08-965-903B-2
Sequence 2, Application US/08965903B
Patent No. 6060275
GENERAL INFORMATION:
APPLICANT: Hachon, NIT
APPLICANT: Krasnow, Mark A.
TITLE OF INVENTION: SPROUTY PROTEIN AND CODING
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Denlinger & Associates
STREET: 350 Cambridge Ave., Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/965, 903B
FILING DATE: 07-NOV-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/030232
FILING DATE: 07-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Petichory, Joanne R
REGISTRATION NUMBER: 42,995
REFERENCE/DOCKET NUMBER: 8600-0177.30
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-324-0880
TELEFAX: 650-324-0960
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 591 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-965-903B-2

Query Match 5.7%; Score 115.5; DB 3; Length 591;
Best Local Similarity 22.4%; Pred. No. 0.013;
Matches 61; Conservative 32; Mismatches 112; Indels 67; Gaps 10;

Qy 144 RRRKOP-----ASTSLASRYGRIAPSLVNGDA-----GNNGFLLNGDEIG-WP 188
Db 24 RRAPEPTLGVDTAGATASALASGASAPVAIHNNNSQOOLSTSAASNNNTISTP 83
Qy 189 SRTLD-----TRVRRPVS-----SPSWQINPMNVFSGYSGRTSFSSPEIMD 234
Db 84 ASPPDDVDQIHLTLPLQRPSSLSRNSSTASTATGIVSGSGSSSFTFTR----- 138
Qy 235 TKLESYKIGIDSCALSLISNPHQPHDNN--NNNNNNNNNNNTWASSGFGPMVTYMAQ 292
Db 139 -----RRPPAPVPLNNSISNNNNNSINNNTFLSHFQSAEPASNALGQ 179
Qy 233 PPAPSQHOYLNPWFKNQNDMSFVNLGRYTEP-----DNCQISSGTAM-----GEF 342
Db 180 PPASVTLAQPRESERLTNEYVDTPLOHATRSQHPAGQDNGQTTTHLLLPORNOHL 239

Qy 343 EISDH--HSGRRQYMEDETRAYDSSSHHTN 372
Db 240 HLCQHCHLQOOOQOOOQOOOQOOOQOOOQOOOQOOOQOOOQOOOQOOO 271

RESULT 3

US-08-317-310A-15
Sequence 15, Application US/08317310A
Patent No. 5858701
GENERAL INFORMATION:
APPLICANT: WHITE, Morris F.
APPLICANT: SUN, Xiao Jian
APPLICANT: PIERCE, Jacalyn H.
TITLE OF INVENTION: THE IRS FAMILY OF GENES
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/317,310A
FILING DATE: 03-OCT-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Louis Myers
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: JDP-022
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 1234 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
US-08-317-310A-15

Query Match 5.1%; Score 105; DB 2; Length 1234;
Best Local Similarity 21.0%; Pred. No. 0.36;
Matches 104; Conservative 57; Mismatches 176; Indels 158; Gaps 25;

Qy 1 MEMGNSGPGHGPQGAEEGSGSTESSPSFGGLMFQKITYPEDG----- 44
Db 385 VLSSTSTSGHG-STSDCLFRRRSASVSGSPSDGFLSSDEYSGSPDDFSSFSFSTVPD 443
Qy 45 -----GSGSGS-----SSGGRSNNRVVGGSGSGGSGGQIQR 73
Db 444 SLGHTPPARGEBELSNVICMGOKGASTLAEPNGHYILSRGGNGHRYIGANLGTSPALPG 503
Qy 74 COVECCGMDLTNAKYRYRHRVCV-----HSTP-KTVVAGIEGRFOQCSRF----- 121
Db 504 DEAGAA-DLNR--FRKRTSAGTSPITISHQKTPSOSSVASIEEYTEMMPAAVPPGGG 560
Qy 122 -----HQPEFDLEKSGCRRLAGHNERRRKQPALEVLASRYGRIAPS 166
Db 561 GGRLPCHYHSAFVPTHYSPEEGLEMHLEER-GGH-----RPDTSNLT-DDGYMPMP- 613
Qy 167 LYENDAGNGSGFLNGOETIGWPSRTLDTRVRRFVSSPSMQIND-----YN 213
Db 614 -----GVAFVPSNRKNGY-WPMS-----PKSVSAPOQIINIRHHPQVDPNGYGM 660
Qy 214 VFSQGS-----VGGGTFSSPEIMDKLESY-----KIGDSNC-ALSLSLNPHO----- 258

Db 661 MSPGSCSPDIGGSSSSSSISAAPSG-SSYCKPWTNGVGHHTHALPHAKPVESGGCK 719
Qy 259 --PHDNNNNNNNNNNNTWRASSGFGPMTVTWAQPPAPSOHQYLNPPWFKDNDNM 316
Db 720 LLPCTGYMMSPVGDSTSSPSECYGP-----EDQHKPVLSTYSLSRPFHGTORG 773
Qy 317 SPV-----LNLG-----RY--TEPDNCOISSGTAMG-----EFELSDHHQSR 353
Db 774 EPEGARHOLRLSSSGRLRYTATAEDSSSTSDSLGGGYCGARPESSLTHPHHVLQ 833
Qy 354 QYMEDENTRAVDSSS 368
Db 834 PHLPKVDYTAQTNS 848

RESULT 4
PCT-US95-13041-15
Sequence 15, Application PC/TUS9513041
GENERAL INFORMATION:
APPLICANT: WHITE, Morris F.
APPLICANT: SUN, Xiao Jian
APPLICANT: PIERCE, Jacalyn H.
TITLE OF INVENTION: THE IRS FAMILY OF GENES
NUMBER OF SEQUENCES: 63
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/13041
FILING DATE: Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/317,310
FILING DATE: 03-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Louis Myers
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: JDP-022PC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 1234 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
PCT-US95-13041-15

Query Match 5.1%; Score 105; DB 5; Length 1234;

Best Local Similarity 21.0%; Pred. No. 0.36; Mismatches 176; Indels 158; Gaps 25;

Matches 104; Conservative 57; Mismatches 176; Indels 158; Gaps 25;
Qy 1 MEMSGNSGPGHGPQAEBSGSGSTSSPSGGLMFGQKLYFEDGG----- 44
Db 385 VSLSSSTSGHG-STSDCLFPRRSSASVSGSPDGGFISDEYSSPCDDPSSFRSSTPD 443
Qy 45 -----GGSGSS-----SSGGRSNRRVAGSGSGGSGQIPR 73
Db 444 SLGHTTPARGEEELSNYICMGKGASTLAAPNGHYILSRGGNGRHYIRGANLGTSPALPG 503
Qy 74 CQVCGGMDLTMAGYIRHAYCV-----HKTLP-KYTVAGIERPQCGSRF----- 121
Db 504 DEAGAA-DLNR--FRKRTSAGTSPTISHQKTPSSQSSVASIEYTEMPPAAYPPGGGS 560

Qy 122 -----HQLPFDLEKRSRRRLAGHNRERRRPOPASLVSRYGRAP 166
Db 561 GGRLPGRHSAPVTHSYPEEGLEMHLERR-CGHH-----RPDTSNLHT-DDGTPMSP 613
Qy 167 LYENGDAWGNGSFLGNOETMPPSSRTLDTRVMRPVSSPQWQNP-----NN 213
Db 614 ---GVAPVPSNRKQNDY-MPMS-----PKSYAPQOIIPIPIRHQRYDPNGY 660
Qy 214 VFSQGS---VGGRTSFSPPEIMDTKLESY-----KGIDSNQ-ALSLSNPHQ----- 258
Db 661 MSPGSCSPDIGGSSSSSSISAAPSG-SSYCKPWTNGVGHHTHALPHAKPVESGGCK 719
Qy 259 --PHDNNNNNNNNNNNTWRASSGFGPMTVTWAQPPAPSOHQYLNPPWFKDNDNM 316
Db 720 LLPCTGYMMSPVGDSTSSPSECYGP-----EDQHKPVLSTYSLSRPFHGTORG 773
Qy 317 SPV-----LNLG-----RY--TEPDNCOISSGTAMG-----EFELSDHHQSR 353
Db 774 EPEGARHOLRLSSSGRLRYTATAEDSSSTSDSLGGGYCGARPESSLTHPHHVLQ 833
Qy 354 QYMEDENTRAVDSSS 368
Db 834 PHLPKVDYTAQTNS 848

RESULT 5
US-08-446-855A-2

Sequence 2, Application US/08446855A
Patent No. 5849573
GENERAL INFORMATION:
APPLICANT: Stewart, Thomas S
APPLICANT: Flores, Maria V
APPLICANT: O'Sullivan, William J
TITLE OF INVENTION: Nucleotide sequence encoding carbamoyl
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon & Vanderhye PC
STREET: 1100 No. 5849573th Glebe Road, 8th Floor
CITY: Arlington
STATE: Virginia
COUNTRY: USA
ZIP: 22201-4714

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/446,855A
FILING DATE: 06-Jul-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mitchard, Leonard C
REGISTRATION NUMBER: 29,009
REFERENCE/DOCKET NUMBER: 47-80
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4000
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2391 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-446-855A-2

Query Match 5.1%; Score 105; DB 2; Length 2391;
Best Local Similarity 26.2%; Pred. No. 0.93; Mismatches 34; Indels 38; Gaps 4;
Matches 34; Conservative 18; Mismatches 34; Indels 38; Gaps 4;

```
Query Match Similarity      5.1%; Score 105; DB 4; Length 2391;  
Best Local Similarity     26.2%; Pred. No. 0.93;  
Matches    34; Conservative   18; Mismatches   40; Indels    38; Gaps       4  
  
QY      242 GTGDSNCALSLISNPHQHNNNNNNNNNNNTWRASSGFGEPTVTMAQPAPASGHQ 301  
DB          :|||::||||  
DB      385 GYYDVAKKNTNISLNKKICEONNNNEKNKKNKNN----- 419  
  
QY      302 YLNPWFVFDNDNMDSPLNLGRYTEPDCIQSISSGTAMGEFELSDHHOSRRQYMDE-- 359  
           ::|||:::-:||-:|-:|:  
DB      420 --EVDYIKDKEDENNNAVKFYSGYN--NNAQNHEH---EFNLNNDYSIYIRKKMNEEF 472  
  
QY      360 ---NTRAVD 365  
DB      473 LNLVNKRKYD 482
```

..

```
RESULT 7  
US-09-451-117-2  
Sequence 2, Application US/09451117  
Patent No. 6277973  
GENERAL INFORMATION:  
APPLICANT: Jenkins, Mark C.  
APPLICANT: Fayer, Ronald  
APPLICANT: Trout, James  
TITLE OF INVENTION: Cloning and Expression of a DNA Sequence Encoding A 41  
Patent No. 6277973  
TITLE OF INVENTION: kds Cryptosporidium parvum Oocyst Wall Protein  
FILE REFERENCE: 0046.99  
CURRENT APPLICATION NUMBER: US/09/451,117  
CURRENT FILING DATE: 1999-11-30  
NUMBER OF SEQ ID NOS: 4
```

```

US-08-861-464-6
Sequence 6, Application US/08861464
Patent No. 5874210
GENERAL INFORMATION:
APPLICANT: Guarente, Leonard P.
APPLICANT: Austriaco Jr., Nicanor
APPLICANT: Kennedy, Brian
TITLE OF INVENTION: Genes Determining Cellular Senescence
TITLE OF INVENTION: In Yeast
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESSES:
ADDRESSER: Hamilton, Brook, Smith & Reynolds, P. C.
STREET: Two Militia Drive
City: Lexington
STATE: MA
COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/861,464
FILING DATE: 22-MAY-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/396,001
FILING DATE: 28-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/09351
FILING DATE: 15-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/107,408
FILING DATE: 16-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: MIT-6408A22
TELECOMMUNICATION INFORMATION:
TELEPHONE: 781-861-6240
TELEFAX: 781-861-9540
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 888 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-861-464-6
Query Match 5.0%; Score 102; DB 2; Length 888;
Best Local Similarity 26.2%; Pred. No. 0.44;

```


REGISTRATION NUMBER: 30,727
 REFERENCE/DOCKET NUMBER: 19603/1741
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (716) 263-1304
 TELEFAX: (716) 263-1600
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 424 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-120-817-2

Query Match 4.9%; Score 100; DB 4; Length 424;
 Best Local Similarity 20.1%; Pred. No. 0.24;

Matches 83; Conservative 49; Mismatches 128; Indels 152; Gaps 20;

QY 6 NSGPHGPGQAEISGSSSTSSPSGIMFGQKIYFEDGSG-----SGSSSSGGRNRRV 60
 DB 117 NNG-GIETPSADSGGGGTPTATGAGG---GDTPSATGGGGGTPTATGAGGSGGGGTPTA 172
 QY 61 RGGGS-----GSGGQIPROVEGCGMDLTNAKYRHRVCGVHSTKPKTV----- 107
 DB 173 TGGGSGGTPTATGAGGSGGTPTATGAGGSGGTPTATGAGGSGGTPTATGAGGSGGTPTA 230
 QY 108 --AGIEORFCQCCSRF-----HOLPEFLEKRSRRRLAGHNERRRKPPQAS 152
 DB 231 IKVAGEVFPFGHATFTADSMGNGDGENOKPMFELAEATLKNV----- 276
 QY 153 LSVLASRYGRHAPSLYENGDAWMNGSFLGNOEIGWPSRTLDTRVMRPPVSSPSWQINPM 212
 DB 277 -----NLGENEVDGIHVAKNAQEV-----TLD-----NVHAQ 304
 QY 213 NVPS-----QSGVGGRTSSSPRIMDTKESYKIGDSKALSLSPHPDNNNNN 267
 DB 305 NVEEDLITVAGEGAATN-----LTKNSAKGADK--VQNLNANTLTKIDNFK-- 353
 QY 268. NNNNNNNTRASSGFGPTVTMAQPPAPSOHQYLNPPWFKNDNDMSPV-LNLGRY- 325
 DB 354 -----ADRFRTWRT-----NGKQFDDMSIELNGIEANHGKFA 387
 QY 326 ---TEPDNCOISSGTAMGEFELSDHHQSRROYMEDENTRAYD---SSSHT 371
 DB 388 LVKSDSDDLKLT---ATGNIAMTDVKH-----AYDKTQASTQHT 422

RESULT 15
 US-08-914-999-8
 ; Sequence 8, Application US/08914999
 ; Patent No. 6346406

GENERAL INFORMATION:

APPLICANT: Kyazanov, Alexey G.
 APPLICANT: Halt, William N.
 TITLE OF INVENTION: ELONGATION FACTOR-2 KINASE (EF-2 KINASE)
 TITLE OF INVENTION: AND METHODS OF USE THEREFOR
 NUMBER OF SEQUENCES: 25
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: David A. Jackson, Esq.
 STREET: 411 Hackensack Ave, Continental Plaza, 4th
 STREET: Floor
 CITY: Hackensack
 STATE: New Jersey
 COUNTRY: USA
 ZIP: 07601

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/914,999

FILING DATE:
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Jackson Esq., David A.
 REGISTRATION NUMBER: 26,742
 REFERENCE/DOCKET NUMBER: 601-1-078
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 201-487-5800
 TELEFAX: 201-343-1684
 INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 732 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 ORIGINAL SOURCE:
 ORGANISM: Dictyostelium discoideum
 US-08-914-999-8

Query Match 4.9%; Score 100; DB 4; Length 732;
 Best Local Similarity 21.5%; Pred. No. 0.52;

Matches 53; Conservative 28; Mismatches 57; Indels 108; Gaps 11;

QY 168 YENGDAWMNGSFLGNOEIGWPSRTLDTRVMRPPVSSPSWQINPMNVFSGSVGGRTSF 227
 DB 242 YNN-----NSGFVSNDRERTPQSFHFT-----YHSHNQLIIDI--QG-VGDHYT-- 285
 QY 228 SPEIMDTKLESYKIG-----DSNCAL 250
 DB 286 -----DPQHTYDGVGFGIGNLGQKFEKFLDTHKCNALICYLINQISINPKSEKSDCGT 339
 QY 251 -----SLSNPHQHPDNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 306
 DB 340 VPRPDLPDPTSERDNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 375
 QY 307 WFKDNDNMSPYLANGRYTEPDNCOISSGTAMGEFELSDHHQSRROYMEDENTRAYDS 366
 DB 376 ---NNNNNNSSISKSL-----VEISSGK---ERRDRSPSPQLRVSN-----DG 414
 QY 367 SSHHTN 372
 DB 415 NTLNTN 420

Search completed: April 10, 2003, 13:54:59
 Job time : 19 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 10, 2003, 13:53:09 ; Search time 20 Seconds
(without alignments)
1146.301 Million cell updates/sec

Title: US-09-839-185-2

Perfect score: 2040
Sequence: 1 MEMGNSGPGHGPQAESEGG.....MEDENTRAYDSSHTNWSL 375Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 248812 seqs, 61136040 residues

Total number of hits satisfying chosen parameters: 248812

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB pep:*
2: /cgn2_6/ptodata/1/pubppaa/PCR_NEW_PUB pep:*
3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB pep:*
4: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB pep:*
5: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB pep:*
6: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB pep:*
7: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB pep:*
8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB pep:*
9: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB pep:*
10: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB pep:*
11: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB pep:*
12: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB pep:*
13: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB pep:*
14: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2040	100.0	375	10 US-09-839-185-2	Sequence 2, Appl1
2	330	16.2	181	10 US-09-839-185-4	Sequence 4, Appl1
3	321.5	15.8	138	9 US-09-819-142-20	Sequence 20, Appl1
4	306	15.0	848	10 US-09-839-185-8	Sequence 8, Appl1
5	303.5	14.9	174	9 US-09-819-142-8	Sequence 8, Appl1
6	292.5	14.3	131	10 US-09-839-185-6	Sequence 6, Appl1
7	121	5.9	2150	9 US-10-135-322-17	Sequence 17, Appl1
8	120.5	5.9	605	9 US-09-741-233A-2	Sequence 2, Appl1
9	118	5.8	961	10 US-09-801-368-132	Sequence 132, Appl1
10	115	5.6	1569	9 US-10-108-605-303	Sequence 303, Appl1
11	112	5.5	758	10 US-09-801-368-224	Sequence 224, Appl1
12	111.5	5.5	666	10 US-09-801-368-36	Sequence 36, Appl1
13	109	5.3	824	10 US-09-801-368-312	Sequence 312, Appl1
14	107	5.2	411	10 US-09-749-728B-1	Sequence 1, Appl1
15	106	5.2	860	9 US-09-781-693A-2	Sequence 2, Appl1
16	103	5.0	283	10 US-09-864-761-36720	Sequence 36720, A
17	103	5.0	732	9 US-09-832-292-12	Sequence 12, Appl1
18	102	5.0	888	10 US-09-826-752-6	Sequence 6, Appl1
19	101.5	5.0	575	9 US-10-108-605-55	Sequence 55, Appl1

20	100.5	4.9	676	10 US-09-801-368-302	Sequence 302, Appl1
21	100.5	4.9	3092	10 US-09-801-368-172	Sequence 172, Appl1
22	100	4.9	424	10 US-09-835-684-9	Sequence 9, Appl1
23	100	4.9	424	10 US-09-880-371-9	Sequence 9, Appl1
24	100	4.9	424	10 US-09-879-248-14	Sequence 14, Appl1
25	100	4.9	732	10 US-09-994-465-8	Sequence 8, Appl1
26	100	4.9	1207	9 US-10-108-605-71	Sequence 71, Appl1
27	99.5	4.9	861	9 US-09-820-843A-109	Sequence 109, Appl1
28	98.5	4.8	357	10 US-09-864-761-35807	Sequence 35807, A
29	98.5	4.8	474	9 US-09-736-457-1812	Sequence 1812, Appl1
30	98.5	4.8	474	9 US-09-802-941-1812	Sequence 1812, Appl1
31	98.5	4.8	474	9 US-09-849-626-1812	Sequence 1812, Appl1
32	98.5	4.8	474	9 US-10-017-754-1812	Sequence 1812, Appl1
33	98.5	4.8	1331	9 US-10-087-464-50	Sequence 50, Appl1
34	98	4.8	623	9 US-10-108-605-125	Sequence 125, Appl1
35	98	4.8	623	9 US-09-880-748-1415	Sequence 129, Appl1
36	97.5	4.8	252	9 US-09-880-748-1415	Sequence 1415, Appl1
37	97.5	4.8	440	9 US-09-932-367A-19	Sequence 19, Appl1
38	97.5	4.8	633	9 US-10-086-464-11	Sequence 11, Appl1
39	97	4.8	127	10 US-09-916-940-43	Sequence 43, Appl1
40	97	4.8	251	9 US-09-880-748-1084	Sequence 1084, Appl1
41	97	4.8	386	9 US-10-050-704-100	Sequence 100, Appl1
42	96.5	4.7	570	9 US-09-847-102A-43	Sequence 43, Appl1
43	96	4.7	724	9 US-10-108-605-307	Sequence 307, Appl1
44	96	4.7	645	10 US-09-919-172-41	Sequence 41, Appl1
45	96	4.7	655	10 US-09-205-658-57	Sequence 57, Appl1

ALIGNMENTS

RESULT 1
US-09-839-185-2
Sequence 2, Appl1 Application US/09839185
Patent No. US20020069433A1
GENERAL INFORMATION:
APPLICANT: Schmidt, Edward Daniel Leendert
De Vries, Sape Cornelis
Hecht, Valerie France Gabrielle
TITLE OF INVENTION: Apomixis Conferred By Expression of SRK
Interacting Proteins
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Syngenta Patent Dept.
STREET: 3054 Cornwallis Road
CITY: RTP
STATE: NC
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/839,185
FILING DATE: 19-Apr-2001
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Meligs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: S-30683A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 375 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO

ORIGINAL SOURCE:
ORGANISM: Arabidopsis thaliana
IMMEDIATE SOURCE:
CLONE: 3A35
SEQUENCE DESCRIPTION: SEQ ID NO: 2
US-09-839-185-2

Query Match 100.0%; Score 2040; DB 10; Length 375;
Best Local Similarity 100.0%; Pred. No. 1e-163;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEMGNSGPGHGPQDAGSGSSSTESSFSGGLMFGQKTYFEDGGGSGSSSGGRSNRRV 60
DB 1 MEMGNSGPGHGPQDAGSGSSSTESSFSGGLMFGQKTYFEDGGGSGSSSGGRSNRRV 60
QY 61 RGGSGGSGQIPRCQVEGGMGLTNAGYSSRRHRCVSKTPKTVAGIEGRFCQCCSR 120
DB 61 RGGSGGSGQIPRCQVEGGMGLTNAGYSSRRHRCVSKTPKTVAGIEGRFCQCCSR 120
QY 121 FHLPEFLEKSCRRRLAGHNERRRKPPASLASYLASHRYGRIAPSLYENGAGMNGSFL 180
DB 121 FHLPEFLEKSCRRRLAGHNERRRKPPASLASYLASHRYGRIAPSLYENGAGMNGSFL 180
QY 181 GNOETGMPSSRTLDRVWRPVSSPQWQINPMNVFSQSGVGCGRTSFSSEPEIMDTKLSY 240
DB 181 GNOETGMPSSRTLDRVWRPVSSPQWQINPMNVFSQSGVGCGRTSFSSEPEIMDTKLSY 240
QY 241 KGIGSNCLSLTSPHOPDNN 300
DB 241 KGIGSNCLSLTSPHOPDNN 300
QY 301 QYLNPPWFKDNDNDMSPLYNLGRYTEPDNCOISSGTAMGEFELSDHHQSRROYMEDN 360
DB 301 QYLNPPWFKDNDNDMSPLYNLGRYTEPDNCOISSGTAMGEFELSDHHQSRROYMEDN 360
QY 361 TRAYSSSHHTWML 375
DB 361 TRAYSSSHHTWML 375

RESULT 2
US-09-839-185-4
Sequence 4, Application US/09839185
Patent No. US2002006943A1
GENERAL INFORMATION:
APPLICANT: Schmidt, Eduard Daniel Leendert
De Vries, Sape Cornelis
Hecht, Valerie France Gabrielle
TITLE OF INVENTION: Apomixis Conferred By Expression of SERK
Interacting Proteins
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Syngenta Patent Dept.
STREET: 3054 Cornwallis Road
CITY: RTP
STATE: NC
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/839,185
FILING DATE: 19-Apr-2001
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: S-30683A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587

INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 181 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Arabidopsis thaliana
IMMEDIATE SOURCE:
CLONE: 3B39
SEQUENCE DESCRIPTION: SEQ ID NO: 4
US-09-839-185-4

Query Match 16.2%; Score 330; DB 10; Length 181;
Best Local Similarity 43.4%; Pred. No. 2.1e-20; Indels 34; Gaps 7;
Matches 75; Conservative 22; Mismatches 42;
QY 45 GSGSSSSGGRSNR-----RVGGSGSGSQIPR--COVEGGMGLTNAGYSSRRHRC 96
DB 29 GMDGEEDGDEDRKKMYRVRG---PSTDVPSRLCQVDRCTVNLTAQYRRHRC 85
QY 97 GVHSTPRVTYAGIEGRFCQCCSRFHQLEPFDLKSCRRRLAGHNERRRKPPASISVL 156
DB 86 EVHAKASATYAGVGRFCQCCSRFHELPEDAKRSCRRRLAGHNERRRK-----IS 138
QY 157 ASRYGRILPSLYENGAGMNGSFLGNOELGWPSSRTLDTRVWR--PVSSPQW 208
DB 139 GDSFGE-----GSGRRGFSQGLIQDTE-----RNRVDRKLPMTNSSFK 176

RESULT 3
US-09-819-142-20
Sequence 20, Application US/09819142
Publication No. US20030041356A1
GENERAL INFORMATION:
APPLICANT: Reuber, Lynne
APPLICANT: Ratcliffe, Oliver
APPLICANT: Jiang, Cai-Zhong
APPLICANT: Riechmann, Jose Luis
APPLICANT: Heard, Jacqueline
FILE REFERENCE: MBI-0033
TITLE OF INVENTION: Methods for Modifying Flowering Phenotypes
CURRENT APPLICATION NUMBER: US/09/819,142
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 20
LENGTH: 138
TYPE: PRT
ORGANISM: Arabidopsis thaliana
US-09-819-142-20

Query Match 15.8%; Score 321.5; DB 9; Length 138;
Best Local Similarity 57.7%; Pred. No. 7.5e-20;
Matches 64; Conservative 11; Mismatches 25; Indels 11; Gaps 3;
QY 45 GSGSSSSGGRSNR-----RVGGSGSGSQIPR--COVEGGMGLTNAGYSSRRHRC 96
DB 29 GMDGEEDGDEDRKKMYRVRG---PSTDVPSRLCQVDRCTVNLTAQYRRHRC 85
QY 97 GVHSTPRVTYAGIEGRFCQCCSRFHQLEPFDLKSCRRRLAGHNERRRK 147
DB 86 EVHAKASATYAGVGRFCQCCSRFHELPEDAKRSCRRRLAGHNERRRK 136

RESULT 4
US-09-839-185-8
Sequence 8, Application US/09839185
Patent No. US2002006943A1


```

GENERAL INFORMATION:
APPLICANT: Schmidt, Eduard Daniel Leendert
De Vries, Sape Cornelis
Hecht, Valerie France Gabrielle
TITLE OF INVENTION: Apomixis Conferred By Expression of SERK
Interacting Proteins
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Syngenta Patent Dept.
STREET: 3054 Cornwallis Road
CITY: RTP
STATE: NC
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/839,185
FILING DATE: 19-Apr-2001
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: S-30683A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 848 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Arabidopsis thaliana
IMMEDIATE SOURCE:
CLONE: 3A52
SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-839-185-8
Query Match 15.0%; Score 306; DB 10; Length 848;
Best Local Similarity 27.9%; Pred. No. 1.6e-17;
Matches 100; Conservative 44; Mismatches 111; Indels 104; Gaps 13;
QY 17 EGGSTSSSSFFSGGLMGCKIYFEDGGGGSSSS-----GGRSNRR---VRGGSSG--- 66
DB 6 DEGG---AAQPFYGSV-----GNSSSSSSSCDEGDKKRAVAIQGDITNGALT 51
QY 67 -----OSGOIPRCQVEGCGMDLTNAKGYSRHYRGVHSTPKVTYAGI 110
DB 52 LNLNGESDGLFPAKTKSGAV--QVENCEADLSKYKDYHRRHKVCHEHSAKATSTVGGI 109
QY 111 EORFQCCSRFHQLEPFLERKSCRRRLAGHNERRRKPOPASLSVLAASYRIADSLYEN 170
DB 110 LQRFQCCSRFHLLEFDFDKRSCRRLAGHNRKRKTP----- 149
QY 171 GDAGNGS-----FLGNOELGMPSSRLDTRVYMRPVSSPSWQINPMNV 214
DB 150 -EPGANGPSDDHSSNYLLITLTKLTSNMH-NHTGDQDLSHLKSLVSHAGEQLKQLV 207
QY 215 -----FSQGSVGGGRTSPSSPE-IMPTKLESYKIGDSNCALSLNPHQPHDNNNNN 266
DB 208 ELLQGRSSQSLNIGNALLGIEQAPQELKQFARQDGTATERRSKQYVMDNFDLND 267
QY 267 NNNNNNNNTWASGFGPMVTYMAQ-----PPASQHOYLNPPVWFKDNNDNDSP 318
DB 268 IYISDDTDVRSPPPTPATPSLDYPSWIHQSSPPOTSRN-----SPSASDOSP 317

```

```

RESULT 5
US-09-819-142-8
Sequence 8, Application US/09819142
Publication No. US20030041356A1
GENERAL INFORMATION:
APPLICANT: Reuber, Lynne
APPLICANT: Ratcliffe, Oliver
APPLICANT: Jiang, Cai-Zhong
APPLICANT: Rieckmann, Jose Luis
APPLICANT: Heard, Jacqueline
APPLICANT: Adam, Luc
TITLE OF INVENTION: Methods for Modifying Flowering Phenotypes
FILE REFERENCE: MBI-0033
CURRENT APPLICATION NUMBER: US/09/819,142
CURRENT FILING DATE: 2001-03-27
NUMBER OF SEQ ID NOS: 24
SOFTWARE: Patentin version 3.0
SEQ ID NO 8
LENGTH: 174
TYPE: PRT
ORGANISM: Arabidopsis thaliana
US-09-819-142-8
Query Match 14.9%; Score 303.5; DB 9; Length 174;
Best Local Similarity 51.2%; Pred. No. 3.3e-18;
Matches 66; Conservative 13; Mismatches 35; Indels 15; Gaps 4;
QY 59 RVRC--GGSGGSGOIPRCQVEGCGMDLTNAKGYSRHYRGVHSTPKVTYAGIQRFCQ 116
DB 37 RVRSRSRISIRGSLRICOVDCTADMKKAKLYHRRHKVCEVAKSSVFLSGINORFCQ 96
QY 117 QCSRFDLPEDLEKSCRRRLAGHNERRRKPOPASLSVLAASYRIADSLYENDGAGNN 176
DB 97 QCSRFDLPEDLEKSCRRRLAGHNERRRK-----SSGESITGE-----GSGRRGIN 144
QY 177 GS-FLNGQE 184
DB 145 GQVVMQNGE 153
RESULT 6
US-09-839-185-6
Sequence 6, Application US/09839185
Patent No. US20020069433A1
GENERAL INFORMATION:
APPLICANT: Schmidt, Eduard Daniel Leendert
De Vries, Sape Cornelis
Hecht, Valerie France Gabrielle
TITLE OF INVENTION: Apomixis Conferred By Expression of SERK
Interacting Proteins
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Syngenta Patent Dept.
STREET: 3054 Cornwallis Road
CITY: RTP
STATE: NC
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/839,185
FILING DATE: 19-Apr-2001
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: S-30683A
TELECOMMUNICATION INFORMATION:

```

RESULT 8

```

; APPLICANT: Summers, Eric
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
; FILE REFERENCE: 109272.147

```



```

; APPLICANT: Sakurada, Kazuhiro
; APPLICANT: Gojo, Satoshi
; APPLICANT: Yamada, Yoichi
; TITLE OF INVENTION: THE CELL HAVING THE POTENTIALITY OF DIFFERENTIATION INTO CARDIOMY
; FILE REFERENCE: 00766.000043
; CURRENT APPLICATION NUMBER: US/09/749,728B
; CURRENT FILING DATE: 2001-09-17
; PRIOR APPLICATION NUMBER: H11-372826
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: PCT-JP00-01148
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: PCT-JP00-07741
; PRIOR FILING DATE: 2000-11-02
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: Patent Ver.2.0
; SEQ ID NO 1
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-749-728B-1

```

```

Query Match          5.2%; Score 107; DB 10; Length 411;
Best Local Similarity 22.8%; Pred. No. 0.33;
Matches 64; Conservative 34; Mismatches 111; Indels 72; Gaps 13;

```

```

QY 4 GSNPGHGPQOASGSGSTSSSFGGLMPGQKIYFEDGGGSGSSSGGRSNR-----59
DB 85 GGRGRGRGRGRPRPSGSGGL-----GG-----DGGCGCGGSGGAGARRRPRPV 128
QY 60 VRGGSGSGSQIIPRCQYEGCGMDLTNAKGYSRHRVCGVSHKTRPVAGIE-----QR 113
DB 129 FPGSGAGPGRGPRATESGKMDCCPALPPGWKKEV-----IRKSGLAGKSDVYY 179
QY 114 FCCGCSRFHQLPEDLEKRSRRRLAGHNERRRKPPASISVLASRYGRIAPSLYENGDA 173
DB 180 FSPSGKKFRSKPOL-----ARYLANT-----VDLSSFDFTGKMPSKLQKNKQ 223
QY 174 GNGSGFLGNOEIGWPS-SRTLDTR-----VMRRPVSS-----NPMNVF 215
DB 224 RLRRDPL-NQNKGRPDINTLPIROTASIFKQPYTKYTNHNSNKVKSDPQMGMEQPRQLF 262
QY 216 SQGSVGGGRTSFSSPEIMDTK-LESYKIGDSNCALSLLS 254
DB 283 WEKRLQGLSASDVTEQIITKIEWLPKGLQGVGPGSNDFTLLS 323

```

```

RESULT 15
US-09-781-693A-2
; Sequence 2, Application US/09781693A
; Publication No. US20030054438A1
; GENERAL INFORMATION:
; APPLICANT: Chang, Tai-Jay
; TITLE OF INVENTION: ANDROGEN RECEPTOR COMPLEX-ASSOCIATED
; FILE REFERENCE: 11709-003001
; CURRENT APPLICATION NUMBER: US/09/781,693A
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US 60/262,312
; PRIOR FILING DATE: 2001-01-17
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 860
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-781-693A-2

```

```

Query Match          5.2%; Score 106; DB 9; Length 860;
Best Local Similarity 19.8%; Pred. No. 1;
Matches 83; Conservative 57; Mismatches 182; Indels 98; Gaps 15;
QY 5 SMSGPGHGP-QOASGSGSTSSSFGGLMPGQKIYFEDGGGSGSSSGGRSNRVRVGG 63

```

```

DB 317 SDTGPRAPRPEREREDGSGPNVSLMQRMSSDMLSRWFEEASEVAQNRGGRPR---G 373
QY 64 GSGQSG-----QIPRCQVEGCGMDL-----TNAKGYSRHRVCGVSH 100
DB 374 GTQSDISTLPTVPSSPDLEVESETAMEVDTPAQFLQPTSSSTMSQAHSSTSSPTSSPHS 433
QY 101 KTRPVTVAGIEORFCQCSRFHQLPEDLEKRSRRRLAGHNERRRKPPAS-LSVL 156
DB 434 -TPLSSPDSEQRQSYEASGHTHQSD-----NNNEKLSPKRGTEGPEVLSLH 480
QY 157 ASRYGRIAPSLYEN-----GDAGMNGSFL-----GNOEIGWP 188
DB 481 YSTEGTTSTIKLFTDEWSSIASSSRGISGSHCKSGQESFVPGSSVOPPEGDSSTKAP 540
QY 189 SSRTLDTRVWRPRVSSPSMQINPMNVFSQGSVCGGRTSFSSPEIMPTKLESYKIGDSNC 248
DB 541 EESSEDTYTKYQEGVSAENPVENHINITQSDKFTAKPLDSNGSRNDLNDRCGVBEESA 600
DB 601 SSEKAKEPETSQDSTESATNENNTNPE-----PQFQTEA---TGFEAHETST--- 646
QY 309 FKD-----NDNMSPYLNLG-RYTEPDNCQISSGTAMGFEISDHHQSRQYMEDENT 361
DB 647 -RDSALQDITDSDDDPYLI.PGARYRAGPGDPRSAAVARIOBF-----FRRRKEREEMELDT 701

```

Search completed: April 10, 2003, 13:55:25
Job time : 25 secs

GenCore version 5.1.4 p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 10, 2003, 13:51:09 ; Search time 35 Seconds
(without alignments)
2207.648 Million cell updates/sec

Title: US-09-839-185-2

Perfect score: 2040

Sequence: 1 MEGNSNGPCHGPGQAEISG.....MEDENTRAYDSSSHHTWISL 375

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2033	99.7	375	10	048527 arabidops
2	2009.5	98.5	374	10	0944J2 arabidops
3	1956	95.9	373	10	09SMX7 arabidops
4	863.5	42.3	354	10	09M206 arabidops
5	615	30.1	323	10	09SNV3 arabidops
6	556	27.3	257	10	09SNV4 arabidops
7	417	20.4	329	10	09SM15 arabidops
8	400	19.6	425	10	09SMY0 arabidops
9	396	19.4	359	10	09LUF4 arabidops
10	392	19.2	419	10	09S840 arabidops
11	374.5	18.4	394	10	094TW8 arabidops
12	374.5	18.4	405	10	09S7A2 arabidops
13	365.5	17.9	378	10	09LOA4 arabidops
14	356	17.5	392	10	09FZJ9 arabidops
15	356	17.5	396	10	09ZWS3 arabidops
16	356	17.5	396	10	08S9L0 arabidops

17	354.5	17.4	389	10	09FZK0 arabidops
18	354.5	17.4	393	10	09ZW54 arabidops
19	351	17.2	367	10	09SM14 arabidops
20	334	16.4	435	10	09SM17 arabidops
21	333	16.3	333	10	09S849 arabidops
22	330	16.2	181	10	09S758 arabidops
23	328.5	16.1	1020	10	09C9K0 arabidops
24	327.5	16.1	424	10	09SM16 arabidops
25	326.5	16.0	440	10	09SM19 arabidops
26	321.5	15.8	138	10	09SMW8 arabidops
27	321	15.7	927	10	09S7P5 arabidops
28	318.5	15.6	1035	10	09S7N9 arabidops
29	313	15.3	305	10	09SNV5 arabidops
30	311.5	15.3	445	10	09SM18 arabidops
31	310	15.2	881	10	082651 arabidops
32	309	15.1	187	10	09SMX9 arabidops
33	303.5	14.9	167	10	09SMX8 arabidops
34	303.5	14.9	174	10	09S7A9 arabidops
35	297	14.6	240	10	080718 arabidops
36	292.5	14.3	129	10	09LFB4 arabidops
37	292.5	14.3	131	10	P93015 arabidops
38	283	13.9	862	10	09LGV7 arabidops
39	272	13.4	131	10	09SM13 arabidops
40	255.5	12.5	377	10	09SMY1 arabidops
41	254.5	12.5	775	10	08VZV0 arabidops
42	254.5	12.5	789	10	08S9G8 arabidops
43	254.5	12.5	801	10	09S7Z3 arabidops
44	169	8.3	662	5	08TIP5 dicysterei
45	160.5	7.9	816	5	024538 drosophila

ALIGNMENTS

RESULT 1
ID 048527 PRELIMINARY; PRT; 375 AA.
AC 048527;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Putative squamosa-promoter binding protein.
GN T24P15.11 OR SPL9.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eudicots II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Rounsley S.D., Lin X., Ketchum K.A., Crosby M.L., Brandon R.C.,
RA Sykes S.M., Kaul S., Mason T.M., Kerlavage A.R., Adams M.D.,
RA Somerville C.R., Venter J.C.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA MEDLINE=97446501; PubMed=9301089;
RA Cardon G.H., Hoehmann S., Nettesheim K., Saedler H., Huisser P.;
RT "Functional analysis of the Arabidopsis thaliana SBP-box gene SPL3: a
RL novel gene involved in the floral transition.";
RN Plant J. 12:367-377 (1997).
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA MEDLINE=99453765; PubMed=10524240;
RA Cardon G.H., Hoehmann S., Klein J., Nettesheim K., Saedler H.,
RA Huisser P.;
RT "Molecular characterization of the Arabidopsis SBP-box genes.";
RL Gene 237:91-104 (1999).
DR EMBL; AC002561; AAB8645.1; -;
DR EMBL; AJ011640; CAB56592.1; -;

DR EMBL: AJ011638; CAB56590.1; -
 DR InterPro: IPR004333; SBF_plant_prot.
 DR Pfam: PF03110; SBF; 1.
 SQ SEQUENCE 375 AA; 40846 MW; D1B93F98D70ECD28 CRC64;

Query Match 99.7%; Score 2033; DB 10; Length 375;
 Best Local Similarity 99.7%; Pred. No. 3.5e-153;
 Matches 374; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MEMGNSGPHGPGQAEBSGSGSTSSSFSGGLMFQKTYFEDGGGSGSSSGRSNRV 60
 DB 1 MEMGNSGPHGPGQAEBSGSGSTSSSFSGGLMFQKTYFEDGGGSGSSSGRSNRV 60
 QY 61 RGGSGGSGQIPRCQVECGMDLTNAKGYSRHRCVGHSTPKYTVAGIEQRFCQCCSR 120
 DB 61 RGGSGGSGQIPRCQVECGMDLTNAKGYSRHRCVGHSTPKYTVAGIEQRFCQCCSR 120
 QY 121 FHQLPEPDLERKSCRRLAGHNERRRKQPASLSVLSRYRIAPSLYENGDAWNQSF 180
 DB 121 FHQLPEPDLERKSCRRLAGHNERRRKQPASLSVLSRYRIAPSLYENGDAWNQSF 180
 QY 121 FHQLPEPDLERKSCRRLAGHNERRRKQPASLSVLSRYRIAPSLYENGDAWNQSF 180
 DB 121 FHQLPEPDLERKSCRRLAGHNERRRKQPASLSVLSRYRIAPSLYENGDAWNQSF 180
 QY 121 GNOEIGWPSRFTLDRVWRPVSSPQWQINPMNVFSQSGVGSGRTSFSSPEIMDTKLESY 240
 DB 121 GNOEIGWPSRFTLDRVWRPVSSPQWQINPMNVFSQSGVGSGRTSFSSPEIMDTKLESY 240
 QY 121 GNOEIGWPSRFTLDRVWRPVSSPQWQINPMNVFSQSGVGSGRTSFSSPEIMDTKLESY 240
 DB 121 GNOEIGWPSRFTLDRVWRPVSSPQWQINPMNVFSQSGVGSGRTSFSSPEIMDTKLESY 240
 QY 241 KGIGDNCALSLISNPHQPHDNNNNNNNNNNNTWRASGFGPMVTYMAQPPAPSOH 300
 DB 241 KGIGDNCALSLISNPHQPHDNNNNNNNNNNNTWRASGFGPMVTYMAQPPAPSOH 300
 QY 301 QYLNPPWVFKDNDMSFVLNIGRYTEPDNCQISSGTAMGEFELSDHHQSRQYMEDN 360
 DB 301 QYLNPPWVFKDNDMSFVLNIGRYTEPDNCQISSGTAMGEFELSDHHQSRQYMEDN 360
 QY 361 TRAYDSSSHHTNWSL 375
 DB 361 TRAYDSSSHHTNWSL 375

RESULT 2

Q94AJ2 PRELIMINARY; PRT; 374 AA.

AC Q94AJ2; 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Putative squamosa-promoter-binding protein.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N. A.
 RA Yamada K., Liu S.-X., Pham P.-K., Banh J., Dale J.M., Gibson H.A.,
 RA Goldsmith A.D., Jiang P.-X., Lee J.M., Onodera C.S., Quach H.L.,
 RA Tang C., Toriumi M., Yu G., Yu S., Bowser L., Carlini P., Chen H.,
 RA Chouk R., Hayashizaki Y., Ishida J., Jones T., Kaniya A.,
 RA Karlin-Neumann G., Kawai J., Kim C., Koeseema E., Lam B., Lin J.,
 RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
 RA Sakurai T., Satou M., Seki M., Shin M., Southwick A., Tracy S.E.,
 RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.,
 RT "Full length cDNA of gene T24P15.11/At2g42200 (GI:2673911)." ;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY046007; AAK76681.1; -
 DR InterPro: IPR004333; SBF_plant_prot.
 DR Pfam: PF03110; SBF; 1.
 SQ SEQUENCE 374 AA; 40733 MW; EA3C1C38200E5386 CRC64;

Query Match 98.5%; Score 2009.5; DB 10; Length 374;
 Best Local Similarity 99.2%; Pred. No. 2.6e-151;
 Matches 372; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
 QY 1 MEMGNSGPHGPGQAEBSGSGSTSSSFSGGLMFQKTYFEDGGGSGSSSGRSNRV 60

DB 1 MEMGNSGPHGPGQAEBSGSGSTSSSFSGGLMFQKTYFEDGGGSGSSSGRSNRV 60
 QY 61 RGGSGGSGQIPRCQVECGMDLTNAKGYSRHRCVGHSTPKYTVAGIEQRFCQCCSR 120
 DB 61 RGGSGGSGQIPRCQVECGMDLTNAKGYSRHRCVGHSTPKYTVAGIEQRFCQCCSR 120
 QY 121 FHQLPEPDLERKSCRRLAGHNERRRKQPASLSVLSRYRIAPSLYENGDAWNQSF 180
 DB 121 FHQLPEPDLERKSCRRLAGHNERRRKQPASLSVLSRYRIAPSLYENGDAWNQSF 180
 QY 121 FHQLPEPDLERKSCRRLAGHNERRRKQPASLSVLSRYRIAPSLYENGDAWNQSF 180
 DB 121 FHQLPEPDLERKSCRRLAGHNERRRKQPASLSVLSRYRIAPSLYENGDAWNQSF 180
 QY 121 GNOEIGWPSRFTLDRVWRPVSSPQWQINPMNVFSQSGVGSGRTSFSSPEIMDTKLESY 240
 DB 121 GNOEIGWPSRFTLDRVWRPVSSPQWQINPMNVFSQSGVGSGRTSFSSPEIMDTKLESY 240
 QY 241 KGIGDNCALSLISNPHQPHDNNNNNNNNNNNTWRASGFGPMVTYMAQPPAPSOH 300
 DB 241 KGIGDNCALSLISNPHQPHDNNNNNNNNNNNTWRASGFGPMVTYMAQPPAPSOH 300
 QY 301 QYLNPPWVFKDNDMSFVLNIGRYTEPDNCQISSGTAMGEFELSDHHQSRQYMEDN 360
 DB 301 QYLNPPWVFKDNDMSFVLNIGRYTEPDNCQISSGTAMGEFELSDHHQSRQYMEDN 360
 QY 361 TRAYDSSSHHTNWSL 375
 DB 361 TRAYDSSSHHTNWSL 374

RESULT 3

Q9SMX7 PRELIMINARY; PRT; 373 AA.

AC Q9SMX7; 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Squamosa promoter binding protein-like 9.
 GN SPL9.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N. A.
 RA Cardon G.H., Hochmann S., Nettesheim K., Saedler H., Huijser P.;
 RC MEDLINE=97446501; PubMed=9301089;
 RX MEDLINE=97446501; PubMed=9301089;
 RA Cardon G.H., Hochmann S., Nettesheim K., Saedler H., Huijser P.;
 RT "Functional analysis of the Arabidopsis thaliana SBF-box gene SPL3: a
 RT novel gene involved in the floral transition." ;
 RL Plant J. 12:367-377(1997).
 RN [2]
 RP SEQUENCE FROM N. A.
 RC STRAIN=CV. LANDSBERG ERECTA; TISSUE=FLOWER;
 RX MEDLINE=99453765; PubMed=10524240;
 RA Cardon G.H., Hochmann S., Klein J., Nettesheim K., Saedler H.,
 RA Huijser P.;
 RT "Molecular characterization of the Arabidopsis SBF-box genes." ;
 RL Gene 237:91-104(1999)
 DR EMBL: AJ011639; CAB56591.1; -
 DR InterPro: IPR004333; SBF_plant_prot.
 DR Pfam: PF03110; SBF; 1.
 SQ SEQUENCE 373 AA; 40552 MW; 3B97A844659B2579 CRC64;

Query Match 95.9%; Score 1956; DB 10; Length 373;
 Best Local Similarity 96.1%; Pred. No. 4.4e-147;
 Matches 366; Conservative 0; Mismatches 1; Indels 14; Gaps 2;
 QY 1 MEMGNSGPHGPGQAEBSGSGSTSSSFSGGLMFQKTYFEDGGGSGSSSGRSNRV 60
 DB 1 MEMGNSGPHGPGQAEBSGSGSTSSSFSGGLMFQKTYFEDGGGSGSSSGRSNRV 60
 QY 61 RGGSGGSGQIPRCQVECGMDLTNAKGYSRHRCVGHSTPKYTVAGIEQRFCQCCSR 120
 DB 61 RGGSGGSGQIPRCQVECGMDLTNAKGYSRHRCVGHSTPKYTVAGIEQRFCQCCSR 120


```

Db 61 RGGSGSGSQIPRCQVCGCMTLNAGYYSRRHVCVGHSTKPTVAGIEORFCQCCSR 120
Qy 121 PHOLPEPDLKRSRRRLAGNERRRKQPASLSVLSARYRIAPSLYENGDAAGMSFL 180
Db 121 PHOLPEPDLKRSRRRLAGNERRRKQPASLSVLSARYRIAPSLYENGDAAGMSFL 180
Qy 181 GNOEIGMPPSRRLDTRVRRPVSPPSQINPMNVFSQGSVGGRTFSFSPIMDTKLESY 240
Db 181 GNOEIGMPPSRRLDTRVRRPVSPPSQINPMNVFSQGSVGGRTFSFSPIMDTKLESY 240
Qy 241 KGIQDSNCALSLSNPQPHD-----NNNNNNNNNNNTWRASSGFGPMVTYMAQPP 294
Db 241 KGIQDSNCALSLSNPQPHDNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 300
Qy 295 PAPSOHQYLNPPVWFKNNDNMSPVNLGRTPEPNCQISSGTAMGEFELSDHHOSRRQ 354
Db 301 PAPSOHQYLNPPVWFKNNDNMSPVNLG-----QISSGTAMGEFELSDHHOSRRQ 352
Qy 355 YMEDENTRAYDSSSHHTNWSL 375
Db 353 YMEDENTRAYDSSSHHTNWSL 373

```

RESULT 4

```

Q9M206 PRELIMINARY; PRT; 354 AA.
ID Q9M206;
AC Q9M206;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Squamosa promoter-binding protein homolog.
GN T10K17.130.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC euroside 11; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_Taxid=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Benes V., Wurmbach E., Drzonek H., Anzorge W., Mewes H.W., Lemcke K.,
RA Meyer K.F.X., Queller F., Salanoubat M.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Xu Arabidopsis sequencing project;
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL132977; CAB67620.1;
DR InterPro; IPR004333; SBP_plant_prot.
DR Pfam; PF03110; SBP; 1.
SQ SEQUENCE 354 AA; 39672 MW; F84ED9DFC69CB85 CRC64;

```

Query Match 42.3%; Score 863.5; DB 10; Length 354;
 Best Local Similarity 52.7%; Pred. No. 1.5e-60;
 Matches 204; Conservative 32; Mismatches 86; Indels 65; Gaps 17;

```

Qy 14 GQASGSGSSTRESSSFGSLMGOKIYFEDGGSGSSSGSRNRARRGGSGSGSQIP 72
Db 8 GQASGSGSSTRESSSFGSLMGOKIYFEDG-----SGSRNRNR--NTWRASSTTA 57
Qy 73 RCQVCGCMTLNAGYYSRRHVCVGHSTKPTVAGIEORFCQCCSRPHOLPEPDLKRS 132
Db 58 RCQVCGCMTLNAGYYSRRHVCVGHSTKPTVAGIEORFCQCCSRPHOLPEPDLKRS 117
Qy 133 SCRRRLAGNERRRKQPASLSVLSARYRIAPSLYENGDAAGMSFLGNOEIGMPPSR 192
Db 118 SCRRRLAGNERRRKQPASLSVLSARYRIAPSLYENGDAAGMSFLGNOEIGMPPSR 175
Qy 193 LQTRVRRPVSPPSQINPMNVFSQGSVGGRTFSFSPIMDTKLESYKIG 244
Db 176 LQTRVRRPVSPPSQINPMNVFSQGSVGGRTFSFSPIMDTKLESYKIG 217
Qy 245 DSNCAISLISNPHQPHDNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 293

```

```

Db 218 DSSCALSLSNSPIPIH-----QQQLQPTNTWRPSSGFDMSISFSDKTYMAQPPIS 271
Qy 294 -PAPSOHQYLNPPVWF-----KNDNDMSPVNLGRTPEPNCQISSGTAMGEFELSDH 348
Db 272 QPPISTHQYLSQTWIYAGKSNHSHMSPV-----SQISEPADQISNGTMMGFELLY-LH 327
Qy 349 HQSRROYMEDENTRAYDSSSHHTNWSL 375
Db 328 HQSRROYMEDENTRAYDSSSHHTNWSL 354

```

RESULT 5

```

Q9SNV3 PRELIMINARY; PRT; 323 AA.
ID Q9SNV3;
AC Q9SNV3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Squamosa promoter binding protein-homologue 5 (fragment).
GN SBPH5.
OS Antirrhinum majus (Garden snapdragon).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasteride 1; Lamiales; Veroniceae; Antirrhinum.
OX NCBI_Taxid=4151;
RN [1]
RP SEQUENCE FROM N.A.
RA Cardon G.H., Hoehmann S., Nettesheim K., Saedler H., Huisjer P.;
RL MEDLINE=97446501; PubMed=9301089;
RT "functional analysis of the Arabidopsis thaliana SBP-box gene SPL3: a
RT novel gene involved in the floral transition.";
RL Plant J. 12:367-377 (1997).
RN [2]
RP SEQUENCE FROM N.A.
RA Cardon G.H., Hoehmann S., Nettesheim K., Saedler H.,
RA Huisjer P.;
RT "Molecular characterization of the Arabidopsis SBP-box genes.";
RL Gene 237:91-104 (1999).
DR EMBL; AU011623; CAB56570.1;
DR InterPro; IPR004333; SBP_plant_prot.
DR Pfam; PF03110; SBP; 1.
FT NON TER 1 1
SQ SEQUENCE 323 AA; 35124 MW; D2587CFB96A38EBF CRC64;

```

Query Match 30.1%; Score 615; DB 10; Length 323;
 Best Local Similarity 40.0%; Pred. No. 6.5e-41;
 Matches 154; Conservative 50; Mismatches 91; Indels 90; Gaps 18;

```

Qy 18 SGGSTSSSFGSLMGOKIYFEDGGSGSSSGSRNRARRGGSGSGSQIPRCQVE 77
Db 2 AAGGAERSLN---GLKRGKRIYFEEARAKKKGSTG-----VRCQVE 41
Qy 78 GCGGDLTNAGYYSRRHVCVGHSTKPTVAGIEORFCQCCSRPHOLPEPDLKRSRR 137
Db 42 GCGGDLTNAGYYSRRHVCVGHSTKPTVAGIEORFCQCCSRPHOLPEPDLKRSRR 101
Qy 138 LAGNERRRKQPASLSVLSARYRIAPSLYENGDAAGMSFLGNOEIGMPPSR 189
Db 102 LAGNERRRKQPASLSVLSARYRIAPSLYENGDAAGMSFLGNOEIGMPPSR 157
Qy 190 SKRLDTRVRRPVSPPSQINPMNVFSQGSVGGRTFSFSPIMDTKLESYKIG 246
Db 158 SKRLDTRVRRPVSPPSQINPMNVFSQGSVGGRTFSFSPIMDTKLESYKIG 209
Qy 247 NCALSLISNPHQPHDNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 304
Db 210 STALSLISN---QSWGSRNSNPLGTNGN-----GP---TIVQISINCATIGFTC 255
Qy 305 PPVWFKN--DN---DMSPVNLGRT-----EPDNCQISSGTAMGEFELSDHHQ 350

```

```

Db 256 PSWGGGNPADNTSHDMPNLTGQFSSNSHNTGEBGVQLSHG----- 301
Oy .351 SRQYMEDENTRAYDSSSHHTNWSL 375
Db 302 ---QFQDDHSHRGYDSSVQDMHMSL 323

RESULT 6
O9SNV4 PRELIMINARY; PRT; 257 AA.
ID O9SNV4
AC O9SNV4
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DE 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Squamosa promoter binding protein-homologue 4 (Fragment).
GN SBH4.
OS Antirrhinum majus (garden snapdragon).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Lamiales; Veronicaceae; Antirrhinum.
OX NCBI_Taxid=4151;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=FLOWER;
RX MEDLINE=97446501; PubMed=9301089;
RA Cardon G.H., Hoehmann S., Nettesheim K., Saedler H., Huijser P.;
RT "Molecular characterization of the Arabidopsis thaliana SBP-box gene SPL3; a
RT novel gene involved in the floral transition.";
RL Plant J. 12:367-377(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=FLOWER;
RX MEDLINE=99453765; PubMed=10524240;
RA Cardon G.H., Hoehmann S., Klein J., Nettesheim K., Saedler H.,
RA Huijser P.;
RT "Molecular characterization of the Arabidopsis SBP-box genes.";
RL Gene 237:91-104(1999).
DR EMBL; AJ011622; CAB56569.1;
DR InterPro; IPR004333; SBP_plant_prot.
DR Pfam; PF03110; SBP; 1.
FT NON_TER 1
SQ SEQUENCE 257 AA; 27692 MW; 39869PEBA547635A CRC64;

Query Match 27.3%; Score 556; DB 10; Length 257;
Best Local Similarity 49.4%; Pred. No. 2.3e-36;
Matches 127; Conservative 37; Mismatches 73; Indels 20; Gaps 8;

Oy 25 SSSFGGLMFGQKITYFEDGGGSSSSGGSNRRYRGSGSQGIPRCQYEGGMDIT 84
Db 4 SSSSINGLINFGRKITYFEN-VGSSGLQSS--PSKGRSGGVVQGGQPPRCQVEGCKIDLS 59
Oy 85 NAKGYVSRHRCVGVSKTPKTVAGIEGRFCQCGSRFHQLEPFDELEKSCRRRLAGHNER 144
Db 60 DAKAYVSRHRCVGVSKTPKTVAGIEGRFCQCGSRFHQLEPFDELEKSCRRRLAGHNER 119
Oy 145 RRPQAPASIVLASRYGRIAPSLYENGDAAGMNGSFLGNQEIWPS-----SRTLDTRV 197
Db 120 RRPQAPASIVLASRYGRIAPSLYENGDAAGMNGSFLGNQEIWPS-----SRTLDTRV 197
Oy 198 MRRPVSSPQWQINPMNVFSGSVGGGRTSFSSPEIMDTLSEYKGI-GDSNCALSLSNP 256
Db 176 LGNQSLWQSNQGNPLPDFLQGSSTSRPSYSGHGV--SSVECFGVGVSDDSGALSLSN- 232
Oy 257 HPHNNNNNNNNNNNNNN 273
Db 233 -QPWGSRTRASSLGLNN 248

RESULT 7
O9SM15 PRELIMINARY; PRT; 329 AA.
ID O9SM15
AC O9SM15
DT 01-MAY-2000 (TReMBLrel. 13, Created)

```

```

DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE SBP-domain protein 5 (Fragment).
GN SBP5.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoidae; Andropogoneae; Zea.
OX NCBI_Taxid=4577;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV_7232; TISSUE=EARLY FEMALE INFLORESCENCE;
RX MEDLINE=97446501; PubMed=9301089;
RA Cardon G.H., Hoehmann S., Nettesheim K., Saedler H., Huijser P.;
RT "Functional analysis of the Arabidopsis thaliana SBP-box gene SPL3; a
RT novel gene involved in the floral transition.";
RL Plant J. 12:367-377(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV_7232; TISSUE=EARLY FEMALE INFLORESCENCE;
RX MEDLINE=99453765; PubMed=10524240;
RA Cardon G.H., Hoehmann S., Klein J., Nettesheim K., Saedler H.,
RA Huijser P.;
RT "Molecular characterization of the Arabidopsis SBP-box genes.";
RL Gene 237:91-104(1999).
DR EMBL; AJ011618; CAB56631.1;
DR InterPro; IPR004333; SBP_plant_prot.
DR Pfam; PF03110; SBP; 1.
FT NON_TER 1
SQ SEQUENCE 329 AA; 34775 MW; CE44BE2583CEFD83 CRC64;

Query Match 20.4%; Score 417; DB 10; Length 329;
Best Local Similarity 35.9%; Pred. No. 3.2e-25;
Matches 127; Conservative 46; Mismatches 97; Indels 84; Gaps 21;

Oy 73 RCQVGGCMTDITNAK-GYVSRHRCVGVSKTPKTVAGIEGRFCQCGSRFHQLEPFDELEK 131
Db 9 RCQVGGCMTDITNAK-GYVSRHRCVGVSKTPKTVAGIEGRFCQCGSRFHQLEPFDELEK 68
Oy 132 RSCRRRLAGHNERRRRPPQAPASIVLASRYGRIAPSLYENGDAAGMNGSFL-----GN 182
Db 69 KSCRRRLAGHNERRRRPPQAPASIVLASRYGRIAPSLYENGDAAGMNGSFL-----GN 182
Oy 183 QEIGWPSRRLDTRVMRRPVSSPSWQ--INPMNVFSGSV-----GGRTSFSSP 230
Db 123 MRGPFVAVRPGGRV-----PGGIQWQAGLDPRH--HQGAVAGYGAHYGSEGGSSSARPP 176
Oy 221 EIMDTKLE---SYKGI-GDSNCALSLSNPHQPHDNNNNNNNNNNNNNTWRASSGFGPM 286
Db 177 VFPGELEPPGGCLAGVPADSSCALSLST--QPWDAAHSHSHA-----APTAGFD- 226
Oy 287 TVTMAQPPAPAS---QHQYLNLP-PWVFKDNDNDMSPVNLGR---YTEPDC--QISSG 336
Db 227 ---GGSVVASLSMAASVITAPSPWTEIDSWGHEG-----GRVPPQLPPDDVPLGEVNSG 277
Oy 337 TA-----MGEFELSDHHQ-----SRQYMEDENTRAYDSSSHHTNWSL 375
Db 278 SSSHGQSGSELELALQGNRAPGSAAPAPARNNGSAGT--PDQAGTMDWSL 329

RESULT 8
O9SMY0 PRELIMINARY; PRT; 425 AA.
ID O9SMY0
AC O9SMY0
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DE 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Squamosa promoter binding protein-like 2.
GN SPL2.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

```

```

OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, LANDSBERG ERECTA; TISSUE=FLOWER;
RX MEDLINE=97446501; PubMed=9301089;
RA Cardon G.H., Hoehmann S., Nettesheim K., Saedler H., Huijser P.;
RT "Functional analysis of the Arabidopsis thaliana SBP-box gene SPL3: a
  novel gene involved in the floral transition.";
RL Plant J. 12:367-377 (1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, LANDSBERG ERECTA; TISSUE=FLOWER;
RX MEDLINE=99453765; PubMed=10524240;
RA Cardon G.H., Hoehmann S., Klein J., Nettesheim K., Saedler H.,
  Huijser P.;
RT "Molecular characterization of the Arabidopsis SBP-box genes.";
RL Gene 237:91-104 (1999).
DR EMBL; AU011625; CAB56577.1; -
DR InterPro; IPR004333; SBP_plant_prot.
DR Pfam; PF03110; SBP; 1.
SQ SEQUENCE 425 AA; 47404 MW; 4D2AE7637BAE404A CRC64;

Query Match          19.6%; Score 400; DB 10; Length 425;
Best Local Similarity 31.0%; Pred. No. 9.9e-24;
Matches 120; Conservative 51; Mismatches 108; Indels 108; Gaps 17;

QY 13 PGQASEG---GSSTSSFSFG---GLMFGQKIYFED-----GGGSGS-----SSSSG 53
DB 97 PGEAKCIDTSPTELESPGSDPVLGLKGRITYFEDFEVENAKGLGLPVSLASSVSP 156
QY 54 GRSNRVRGGSGSGSQIPRCQVEGCGMDLTNAKYYSRHRVCGVSHKTPKVTVAIEOR 113
DB 157 VKXSKSI-----PRLQTPHQVEGCVLULSSADYHKKRHCENHSGKPRVAVSGYER 211
QY 114 FCGCCSFHQLPEFDLEKSCRRRLAGHNERRRKPPA-----SLSLASRGRAPLS 166
DB 212 FCCQCSFHCISEFDEKRSCKRRLSDHNAARRKPNRTYDGPQVDYFVNRFPALJHPR 271
QY 167 LYENGDAAGNAGFLGNQELGWPSSRTLDTRWMRRPVSPQIMPMVFGSGSGV---G 223
DB 272 SEB-----KFL-----WSSKVPVSRVL---WPPALTEISNKLFEHSHFGLDP 314
QY 224 RTSFSPEIMDTK---LESYKGIQ-DSNCALSLSNPHQPHDNNNNNNNNNNNTWRA 279
DB 315 KTSARAEALPSKEKVTISSHWGAQDLDGALSLSS-----NSTTW-V 355
QY 280 SSGGEMTVTMAOPPPAPSOHOYINPPWVFKDNDSVPLN-----LGRYTEPNC 331
DB 356 SSSDQPRRFTLDHNP-----SSNLQPIANRSAAQLSVSGYQWDPDP 397
QY 332 QISSGTAMGEFELSDHHQSRQYMED 358
DB 398 AVEGPTAL-----HNGAGQFNEN 416

```

RESULT 9

```

Q9LUF4 PRELIMINARY; PRT; 359 AA.
AC Q9LUF4;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Similarity to squamosa promoter binding protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC euroside II; Brassicales; Brassicaceae; Arabidopsie.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA; PubMed=10718197;
RX MEDLINE=20181125; PubMed=10718197;
RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,

```

```

RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
  features of the regions of 3,076,755 bp covered by sixty P1 and TAC
  RT clones.";
RL DNA Res. 7:31-63 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RA Kaneko T., Katoh T., Asamizu E., Sato S., Nakamura Y., Kotani H.,
  Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. XI.";
RL Submitted (APR-1999) to the EMBL/Genbank/DBD databases.
DR EMBL; AB023037; BAA96980.1; -
DR EMBL; AB025619; BAB09142.1; -
DR InterPro; IPR004333; SBP_plant_prot.
DR Pfam; PF03110; SBP; 1.
SQ SEQUENCE 359 AA; 39108 MW; D80B89FD03D349F CRC64;

Query Match          19.4%; Score 396; DB 10; Length 359;
Best Local Similarity 34.7%; Pred. No. 1.7e-23;
Matches 115; Conservative 31; Mismatches 77; Indels 108; Gaps 15;

QY 4 GSNSGEGHGP---QASGSGSTSSFSGLMFGQKIYFEDGGSGSSSS----- 52
DB 14 GFQDEPDLSPMDSISFGSGSGKADFSFDLKLGRNI-----GNSSVFEGDTQYI 64
QY 53 -----GGR---SNRRVRGGSGSGSQIPRCQVEGCGMDLTNAKYYSRHRVC 96
DB 65 SLSKMDALAKEGSRSSSKTRNGVG-TMOPICLVDCSDPSNCRERHKKHKVC 123
QY 97 GVASKTPKVTVAIEORFCCQCSRFHQLPEFDLEKSCRRRLAGHNERRRKPPAPASLSVL 156
DB 124 DVASKTPVVTINHGKQFCQCCSFHALFEFDKRSCKRRLDGHNRKRPQPEHI--- 180
QY 157 ASRYGRAPLLENGDAAGNAGS---FLGNQELGWPSSRTLDTRWMRRPVSPSSW---QIN 210
DB 181 -----GRPA-----NFTFGGSKLLEFGSGSHV-FPTSVL-----NDSWNSLV 220
QY 211 PMNVFGSGSVGGGRTP-----SSPEIMDTKLSY----- 240
DB 221 SVAVANAGSSYGSGSQSVVSSSPAKTGIMFPISSPSTSTAKQPFLOEBSSTASL 280
QY 241 -----KGIQDGNALSLSN-----PH--QP 259
DB 281 CERMTSCIHSDCALSLSSSSSVPHLQ 311

```

RESULT 10

```

Q9S840 PRELIMINARY; PRT; 419 AA.
AC Q9S840;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Squamosa promoter binding protein-like 2.
GN SPL2.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC euroside II; Brassicales; Brassicaceae; Arabidopsie.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, COLUMBIA; TISSUE=INFLORESCENCE;
RX MEDLINE=97446501; PubMed=9301089;
RA Cardon G.H., Hoehmann S., Nettesheim K., Saedler H., Huijser P.;
RT "Functional analysis of the Arabidopsis thaliana SBP-box gene SPL3: a
  novel gene involved in the floral transition.";
RL Plant J. 12:367-377 (1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, COLUMBIA; TISSUE=INFLORESCENCE;
RX MEDLINE=99453765; PubMed=10524240;

```

RA Cardon G.H., Hoehmann S., Klein J., Nettesheim K., Saedler H.,
 RA Huijser P.;
 RT "Molecular characterization of the Arabidopsis SBP-box genes.";
 RL Gene 237:91-104(1999).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=COLUMBIA;
 RX MEDLINE=99397451; PubMed=10470850;
 RA Kaneo T., Katoh T., Sato S., Nakamura Y., Asamizu E., Kotani H.,
 RA Miyajima N., Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. IX.
 RT Sequence features of the regions of 1,011,550 bp covered by seventeen
 RT P1 and TAC clones.";
 RL DNA Res. 6:183-195(1999).
 DR EMBL; AJ011626; CAB56578.1; -
 DR EMBL; AJ011624; CAB56576.1; -
 DR EMBL; AB017070; BAB10590.1; -
 DR InterPro; IPR004333; SBP_plant_prot.
 DR Pfam; PF03110; SBP; 1.
 SQ SEQUENCE 419 AA; 46660 MW; 092C3A1DF034EBFF CRC64;

Query Match 19.2%; Score 392; DB 10; Length 419;
 Best Local Similarity 31.5%; Pred. No. 4.2e-23;
 Matches 116; Conservative 49; Mismatches 111; Indels 92; Gaps 16;

QY 13 PQAESG---GSSTSSSSFG---GLMFGQKTFED-----GGGSG-----SSSG 53
 Db 94 PQEPAKIDTSPSLSTLSCGSDPVLGLKLGKRTFEDFWEVENAKGLGLPVLASSVSP 153
 QY 54 GRSNRVVGSGSGGQIPRCVCEGMDLTAKGYSRHRCVGHSTKPTKTVAGIEOR 113
 Db 154 VKKSKI-----PQLQTPHCQVECCNLDLSSAKDYHRKRICENHSKPKVVGSEVR 208
 QY 114 FCOQCSRFHQLPEFLEKRSRRRLAGHNERRRKPOPA-----SLAVLASRYRIAPS 166
 Db 209 FCOQCSRFHCLSEFPEKRSRRRLSDHARRKRPRTTYGKQYDFVNRFLIHR 268
 QY 167 LYENDAGMNGSFLNQEIGWPSRTLDTRY--MRPVSS--PSWQINPMWVFSQSGVGG 223
 Db 269 -----SEKRFIMPSSKHVPSRVLMPOAKTEISDTENHNRGLDPKTKAR 314
 QY 224 RTSFSPEIMDTKLESYKIG--DSNCALSLSLNPHQPHDNNNNNNNNNNNTARRSSG 282
 Db 315 AELFEKEKV--TISHHMGASQDLGALSLS-----NSTTV--VSSS 352
 QY 283 FGPMVTMAQPPPA---PSOHQYLNPPWFKNDNDMSVNLGRYTEPDNCQISSGTAM 339
 Db 353 DQPRFTLDHSSNULQPVNR-----SAAQLNSVSGVMQDPDPAVAGPTRL 399
 QY 340 -----GEF 342
 Db 400 HRNGVGQF 407

RESULT 11

Q94JW8 PRELIMINARY; PRT; 394 AA.

AC Q94JW8;
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Similar to squamosa-promoter binding protein 1 isoform
 DE gi11707009.
 GN F4N2.13.
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
 CC NCB1_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lam B., Southwick A., Karlin-Neumann G., Nguyen M., Miranda M.,
 RA Palm C.J., Bowser L., Jones T., Banh J., Chen H., Cheuk R.,

RA Chung M.K., Kim C., Lin J., Liu S.X., Pham P.K., Sakano H., Shinn P.,
 RA Yamada K., Ecker J., Theologis A., Davis R.W.;
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF370612; AAK43931.1; -
 DR InterPro; IPR004333; SBP_plant_prot.
 DR Pfam; PF03110; SBP; 1.
 SQ SEQUENCE 394 AA; 44636 MW; 9ABFC7FFB7DED9 CRC64;

Query Match 18.4%; Score 374.5; DB 10; Length 394;
 Best Local Similarity 32.4%; Pred. No. 9.5e-22;
 Matches 122; Conservative 38; Mismatches 139; Indels 77; Gaps 14;

QY 15 QAEBSGSSSTSSSGLMFGQKTFE-----DGGGSGSSSSGGRNRRVGGSGSQSG 69
 Db 51 EEBEENISSSSKFSQEL--NRIDFKRSFLDGNDDDTSSGFALPSKSRASNICS 108
 QY 70 QIPRCVCEGMDLTAKGYSRHRCVGHSTKPTKTVAGIEGFCOQCSRFHQLPEFDL 129
 Db 109 QNPICQVYGCSKDLSSSKDYHKKRVCENHSTISVIVNGLEQRFCCQCSRFHLSERDD 168
 QY 130 EKRSRRRLAGHNERRRKPPAPASLVSARYRIAPSLYENDAGMNGSFLNQEIGWPS 189
 Db 169 GKRSRRRLAGHNERRRK--PAFYFLPKRH-----KLTFTSQDVGNKFLNSSLVLP 221
 QY 190 S-----RTLDTRYMR--RPVS--PSWQINPMWVFSQSGVGGRTSPSPETIMTKL 237
 Db 222 SFPSSLYRVIVDEDHRTSRVLSFKDEPTCSMFPTN--EONS--SRYESKPAIYSTEV 276
 QY 238 ESYKIGD-----SNCALSLSLNPHQPHDNNNNNNNNNNNTARRSSGFG 284
 Db 277 SINDLHETASRSTRALSLSAQSQHLKFPNTTSITPQNLHSSST----- 328
 QY 285 PMTYTMAQPPPAPOHOYLNPPWFKNDNDMSVNLGRYTEPDNCQISSGTAMEFEL 344
 Db 329 -----DYHQMEQLWIDPKTNSAG-----SSCKGKSTVDLLQL 365
 QY 345 SDH--HHQSRQYMED 358
 Db 366 SSHLQRIEQQRNYTGD 381

RESULT 12

Q9S7A2 PRELIMINARY; PRT; 405 AA.

AC Q9S7A2;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Squamosa promoter binding protein-like 6.
 GN SPL6 OR F23010.24.
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
 CC NCB1_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RX MEDLINE=97446501; PubMed=9301089;
 RA Cardon G.H., Hoehmann S., Nettesheim K., Saedler H., Huijser P.;
 RT "Functional analysis of the Arabidopsis thaliana SBP-box gene SPL3: a
 RT novel gene involved in the floral transition.";
 RL Plant J. 12:367-377(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RX MEDLINE=99453765; PubMed=10524240;
 RA Cardon G.H., Hoehmann S., Klein J., Nettesheim K., Saedler H.,
 RA Huijser P.;
 RT "Molecular characterization of the Arabidopsis SBP-box genes.";
 RL Gene 237:91-104(1999).
 RN [3]
 RP SEQUENCE FROM N.A.

GenCore version 5.1.4.P5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 10, 2003, 13:50:09 ; Search time 13 Seconds

(without alignments)
1196.433 Million cell updates/sec

Title: US-09-839-185-2

Perfect score: 2040
Sequence: 1 MEMSGNSGPHGPGQAESESG.....MEDENTRAVDSSSHHTWLSL 375Scoring table: BLOSUM62
Gapop 10.0 , Gapept 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt 40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	330	16.2	171	1	SBP2_ANTMA
2	308.5	15.1	399	1	LG1_MAIZE
3	305	15.0	131	1	SBP1_ANTMA
4	160.5	7.9	1043	1	FTP1_DROME
5	157.5	7.7	331	1	COL6_ARATH
6	152	7.5	1584	1	KYK1_DICDI
7	146	7.2	1858	1	P3K2_DICDI
8	125	6.1	647	1	KNRL_DROME
9	125	6.1	865	1	E78A_DROME
10	124	6.1	1230	1	ST20_CANAL
11	123.5	6.1	758	1	CC27_YEAST
12	123.5	6.1	779	1	SRP_DROME
13	123.5	6.1	843	1	MNB_DROME
14	123	6.0	988	1	OMB_DROME
15	122.5	6.0	1195	1	YK6_YEAST
16	122	6.0	1585	1	P3K3_DICDI
17	119.5	5.9	472	1	N49_YEAST
18	119.5	5.9	1131	1	YAB9_YEAST
19	117.5	5.8	971	1	CLA4_CANAL
20	116	5.7	355	1	HKL6_LYCES
21	115.5	5.7	374	1	DMT1_MOUSE
22	115	5.6	448	1	AAC2_DICDI
23	114.5	5.6	589	1	SPY_DROME
24	114.5	5.6	2038	1	PSH_DROME
25	113.5	5.6	316	1	LORI_HUMAN
26	112.5	5.5	490	1	MOT3_YEAST
27	112	5.5	614	1	NRD1_HUMAN
28	111.5	5.5	758	1	YM38_YEAST
29	111.5	5.5	666	1	YEAT_YEAST
30	111.5	5.5	989	1	PRP3_DICDI
31	111.5	5.5	2248	1	CYA1_DROME
32	111	5.4	713	1	YHCA_YEAST
33	111	5.4	1596	1	MAM_DROME

34	110.5	5.4	1905	1	TAGE_DICDI
35	110	5.4	964	1	YIN0_YEAST
36	110	5.4	1341	1	YL78_YEAST
37	109.5	5.4	593	1	K1CJ_HUMAN
38	109.5	5.4	914	1	AZP1_YEAST
39	109	5.3	362	1	AVEN_HUMAN
40	109	5.3	576	1	DEAF_DROME
41	109	5.3	823	1	SCN9_YEAST
42	108.5	5.3	481	1	LORI_MOUSE
43	108.5	5.3	490	1	CAR3_DICDI
44	108.5	5.3	708	1	GBF_DICDI
45	108.5	5.3	850	1	D7_DICDI

ALIGNMENTS

RESULT 1	SBP2_ANTMA	STANDARD;	PRT;	171 AA.
AC	Q38740;			
DT	15-JUL-1999 (Rel. 38, Created)			
DT	15-JUL-1999 (Rel. 38, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	Squamosa-promoter binding protein 2.			
GN	SBP2.			
OS	Antirrhinum majus (Garden snapdragon).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;			
OC	Asteridae; euasterids I; Lamiales; Veroniceae; Antirrhinum.			
OX	NCBI_TaxID=4151;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=cv. Snowman;			
RX	MEDLINE=96158840; PubMed=8569690;			
RA	Klein J., Saedler H., Huisjer P.;			
RT	"A new family of DNA binding proteins includes putative			
RT	transcriptional regulators of the Antirrhinum majus floral meristem			
RT	identity gene SQUAMOSA."			
RL	Mol. Gen. Genet. 250:7-16(1996).			
CC	-1- FUNCTION: PROBABLE TRANSCRIPTIONAL FACTOR. BINDS TO THE PROMOTER			
CC	OF THE SQUAMOSA GENE.			
CC	-1- SUBCELLULAR LOCATION: Nuclear.			
CC	-1- SIMILARITY: CONTAINS 1 SBP DOMAIN.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation-			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@sb-sib.ch).			
CC	-----			
DR	EMBL; X92079; CAA63061.1; -			
DR	Interpro; IPR004333; SBP_plant_prot.			
DR	Pfam; PF03110; SBP; 1.			
KW	Transcription regulation; DNA-binding; Nuclear protein.			
FT	DOMAIN 27 39			
FT	DOMAIN 54 57			
FT	DOMAIN 85 159			
FT	DOMAIN 171 AA; 19466 MW; E8B6B49415755EF CRC64;			
SQ	SEQUENCE			
Query Match	16.2%; Score 330; DB 1; Length 171;			
Best Local Similarity	63.7%; Pred. No. 4.4e-17;			
Matches	65; Conservative 5; Mismatches 28; Indels 4; Gaps 2;			
Qy	46 GSGSSSGSGSNRRVRGSGSGSQIPRCQVEGCGMDLTNAKGYSHRRVGVHSKTPKV 105			
Db	61 GEG---SGSGSGEKGHTASGGGVNAQ-PCCIVENGADLRNCKKYYGRHRCVAVAKAPV 116			
Qy	106 TVAGIEQPFQCCGRFQQLPPEFLDKRSRRRLAGNERRRK 147			
Db	117 SVEGIMQRFQCCGRFQQLPPEFLDKRSRRRLAGNERRRK 158			

```

RESULT 2
ID LG1_MAIZE STANDARD: PRT; 399 AA.
AC 004003;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE LIGULELESS1 protein.
GN LG1.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoidae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. W22; TISSUE=leaf;
RX MEDLINE=97230364; PubMed=9119226;
RA Moreno M.A., Harper L.C., Krueger R.W., Dellaporta S.L., Freeling M.;
RT "Liguleless encodes a nuclear-localized protein required for
RT induction of ligules and auricles during maize leaf organogenesis.";
RL Genes Dev. 11:616-628 (1997).
CC -1- FUNCTION: INVOLVED IN THE FORMATION OF LIGULES AND AURICLES DURING
CC -1- LEAF ORGANOGENESIS.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: LEAF LIGULAR REGION, BLADE AND SHEATH.
CC -1- DEVELOPMENTAL STAGE: DEVELOPING LEAVES. PRESENT AT HIGHEST
CC LEVELS IN VERY YOUNG LIGULES AND AURICLES. PRESENT PRE- AND POST-
CC -1- LIGULE AND AURICLE DEVELOPMENT.
CC -1- SIMILARITY: CONTAINS 1 SBP DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U89496; AB51071.1;
DR InterPro; IPR004333; SBP_plant_prot.
KW Nuclear protein; Developmental protein.
FT DOMAIN 243 259 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 185 260 SBP
FT SEQUENCE 399 AA; 43367 MW; 1B89F304232D7B8C CRC64;
SQ
Query Match 15.1%; Score 308.5; DB 1; Length 399;
Best Local Similarity 31.4%; Pred. No. 3.9e-15;
Matches 86; Conservative 32; Mismatches 91; Indels 65; Gaps 9;
OY 21 STBS--SSFG--GLMGQKIYEDGGGSSSSGGRNRRVGGSGSGQIPRCQ 75
DB 146 SAEENMASFAATOLGLNLYRTYFPPRGYT-----XGHPPRQC 186
OY 76 VEGCGMDLTNAQYYSRHRVGVHSKTP-KYTVAGIEORFCQCGSRFHQLEPDLKSC 134
DB 187 AEGCADLSSARRYRRKRVCEHSHKAPVVTAGLHQFCQCSRFHLDFDDAKSC 246
OY 135 RRRLAGHNRERRRPOP-----ASLSVLASRYGRIAPSLYENDAGMNGSFLGN 182
DB 247 RRRLLDHRNRKRKSKPSDADGDKRAHANKAAAKDKAKSSSKMDIGD-GLGAQILGS 305
OY 183 QEIGWPSSTLTDRVMRRVSSPSSWOIN-----PAMVFSQGSVGGGTSSTSPINDT 235
DB 306 ALISKEOQITMDLGEVVKAVDPKGRASWQOHYGFPHSSSAGSCFPQTQAVSS----- 359
OY 236 KLESYKIGDSNCALSLSNP-----HOPHNNN 264
DB 360 -----DTTSNIGQVOEPLGLFHHQHSHN 384

```

```

RESULT 3
ID SBP1_ANTMA STANDARD: PRT; 131 AA.
AC Q387L;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Squamosa-promoter binding protein 1.
GN SBP1.
OS Antirrhinum majus (Garden snapdragon).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; eunastoids I; Lamiales; Veronicaceae; Antirrhinum.
OX NCBI_TaxID=4151;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Snowman;
RX MEDLINE=96158840; PubMed=8565690;
RA Klein J., Saeedler H., Huijser P.;
RT "A new family of DNA binding proteins includes putative
RT transcriptional regulators of the Antirrhinum majus floral meristem
RT identity gene SQUAMOSA.";
RL Mol. Gen. Genet. 250:7-16(1996).
CC -1- FUNCTION: PROBABLE TRANSCRIPTIONAL FACTOR. BINDS TO THE PROMOTER
CC OF THE SQUAMOSA GENE.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: CONTAINS 1 SBP DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X92369; CA63113.1;
DR InterPro; IPR004333; SBP_plant_prot.
DR Pfam; PF03110; SBP; 1.
KW Transcription regulation; DNA-binding; Nuclear protein.
FT DOMAIN 52 126 SBP
FT SEQUENCE 131 AA; 14919 MW; 71DA57E150E4B23C CRC64;
SQ
Query Match 15.0%; Score 305; DB 1; Length 131;
Best Local Similarity 54.5%; Pred. No. 2e-15;
Matches 60; Conservative 12; Mismatches 28; Indels 10; Gaps 3;
OY 48 GSSSSG-----GSRNRRVRC--GSGGSGQIPR-CVEGCGMDLTNAQYYSRHRVCG 97
DB 16 GSGEGEEDDIDGESSKKTALTPSGKRASTGTSRSCVENCAMETNAKPYHRRKVC 75
OY 98 VHSKTPKTVAGIEORFCQCGSRFHQLEPDLKRSCKRRLAGHNRERK 147
DB 76 FHAKAPVVLHSGLOORFCQCSRFHELSEPDARSGCKRRLAGHNRERK 125

```

```

RESULT 4
ID FTFL_DROME STANDARD: PRT; 1043 AA.
AC P33244;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Nuclear hormone receptor FTZ-F1 (FTZ-F1 alpha).
GN FTZ-F1 OR NR5A3.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

```


RX MEDLINE=91227912; PubMed=1709303;
 RA Latorgna G., Ueda H., Clos J., Wu C.;
 RT "FTZ-F1, a steroid hormone receptor-like protein implicated in the
 RT activation of fushi tarazu.";
 RL Science 252:848-851(1991).
 RN [2]
 RP CHARACTERIZATION, AND SUBUNITS.
 RX MEDLINE=94217714; PubMed=8164672;
 RA Ohno C.K., Ueda H., Pectkovich M.;
 RT "The Drosophila nuclear receptors FTZ-F1 alpha and FTZ-F1 beta
 RT compete as monomers for binding to a site in the fushi tarazu gene.";
 RL Mol. Cell. Biol. 14:3166-3175(1994).
 RN [3]
 RP CHARACTERIZATION, AND MUTAGENESIS.
 RX MEDLINE=93078768; PubMed=1448096;
 RA Ueda H., Sun G.-C., Murata T., Hirose S.;
 RT "A novel DNA-binding motif abuts the zinc finger domain of insect
 RT nuclear hormone receptor FTZ-F1 and mouse embryonal long terminal
 RT repeat-binding protein.";
 RL Mol. Cell. Biol. 12:5667-5672(1992).
 RN [4]
 RP CHARACTERIZATION.
 RX MEDLINE=97172279; PubMed=9020363;
 RA Guichet A., Copeland J.W.R., Erideli M., Hlousek D., Zavorazky P.,
 RA Ho J., Brown S., Percival-Smith A., Krause H.M., Ephrussi A.;
 RT "The nuclear receptor homologue Ftz-F1 and the homeodomain protein
 RT Ftz are mutually dependent cofactors.";
 RL Nature 385:548-552(1997).
 RN [5]
 RP CHARACTERIZATION.
 RX MEDLINE=97172280; PubMed=9020364;
 RA Yu Y., Li W., Su K., Yusea M., Han W., Perrimon N., Pick L.;
 RT "The nuclear hormone receptor Ftz-F1 is a cofactor for the Drosophila
 RT homeodomain protein Ftz.";
 RL Nature 385:552-555(1997).
 CC - FUNCTION: ACTS AS A COFACTOR TO FUSHI TARAZU (FTZ). FACILITATES
 CC THE BINDING OF FTZ TO DNA. BINDS THE SEQUENCE ELEMENT 5'-
 CC YCYGCGCR-3' IN THE ZEBRA ELEMENT OF FTZ. PROBABLY ALSO FUNCTION
 CC AS A RECEPTOR FOR A YET UNKNOWN LIGAND.
 CC - SUBUNIT: MONOMER; FORMS A COMPLEX WITH FTZ.
 CC - SUBCELLULAR LOCATION: Nuclear.
 CC - TISSUE SPECIFICITY: EXPRESSION IN THE PARASEGMENTAL PRIMORDIA
 CC OF THE EMBRYONIC BLASTODERM.
 CC - DEVELOPMENTAL STAGE: FIRST APPEARS IN BLASTODERM EMBRYOS. IT IS
 CC ABSENT IN SUBSEQUENT EMBRYO STAGES, AND THEN REAPPEARS IN LATE
 CC EMBRYOGENESIS TO BE FOUND IN LARVAE, PUPAE AND ADULTS.
 CC - SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
 CC NRS SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
 CC or send an email to license@1sb-sib.ch).
 CC -----
 CC EMBL: M63711; AAA28542.1; -.
 CC HSSP: P19793; 2NUL.
 CC TRANSFAC: T00296; -.
 CC FlyBase: FBgn001078; ftz-fl.
 CC InterPro: IPR000536; Hormone_rec_1lg.
 CC InterPro: IPR001628; ZnF_C4steroid.
 CC Pfam: PF00104; hormone_rec_1.
 CC Pfam: PF00105; zf-C4; 1.
 CC PRINTS: PR00047; STROIDFINGER.
 CC ProDom: PD000035; ZnF_C4steroid; 1.
 CC SMART: SM00430; HOL1; 1.
 CC SMART: SM00399; ZnF_C4; 1.
 CC PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
 CC Receptor; Transcription regulation; DNA-binding; Nuclear protein;
 KW Zinc-finger; Activator.
 FT DOMAIN 29 47 GLN-RICH (OPA-REPEAT).

FT DOMAIN 157 171 POLY-ASN.
 FT DOMAIN 361 367 POLY-GLN (OPA-REPEAT).
 FT DOMAIN 414 497 GLY-RICH.
 FT DNA_BIND 510 575 NUCLEAR RECEPTOR-TYPE.
 FT ZN_FING 510 530 C4-TYPE.
 FT ZN_FING 546 570 C4-TYPE.
 FT DOMAIN 688 740 GLY-RICH.
 FT DOMAIN 837 879 LIGAND-BINDING (POTENTIAL).
 FT MUTAGEN 565 565 C->S: 100-FOLD LESS BINDING.
 FT MUTAGEN 581 581 R->Q: 100-FOLD LESS BINDING.
 FT MUTAGEN 584 584 R->Q: 100-FOLD LESS BINDING.
 FT MUTAGEN 585 585 M->I: 10-FOLD LESS BINDING.
 FT MUTAGEN 586 586 R->Q: 10-FOLD LESS BINDING.
 FT MUTAGEN 587 588 GG->AA: NO BINDING.
 FT MUTAGEN 589 589 R->Q: 10-FOLD LESS REDUCED BINDING.
 FT MUTAGEN 591 591 K->Q: NO EFFECT ON BINDING.
 SQ SEQUENCE 1043 AA; 110234 MW; 0C4539F94921617 CRC64;
 Query Match 7.9%; Score 160.5; DB 1; length 1043;
 Best Local Similarity 24.7%; Pred. No. 0.0004;
 Matches 91; Conservative 41; Mismatches 97; Indels 139; Gaps 23;
 QY 5 SNGSGHGRG---QAESGSGST-----ESSSFGS-----GLMFGQKIYFEDGGCG 46
 DB 413 SGGGGGGGGNNLNGSSGSSAGCGYMLLPQAASSGNNGNPNAGHWSGSV--GNGSGG 470
 QY 47 SGSSSSGGRNRRVRGGGS---GSGGQI-----PRCQVEGCG-----MDLT 84
 DB 471 AGNGAGAGNSGPNPMPGKTATPGHGEVVDFFKLPEELCPV--CGDKVSGHYGLLTCE 528
 QY 85 NAKGYYSR-----HRYGVHSKTPKYTVAGIEGRFCQCGSRFHQLEPFDELEKR 132
 DB 529 SCKGPFRTYQNKKYTCVAFERCHI-DKT-----QKRCPPYC-RFQCKLEVGMKLE 578
 QY 133 SCR--RLAGHN-----ERRRKP-----PALSIVLASRG-----RIAPSLYENG 171
 DB 579 AVRRDRRRGRGRNRPYKRRKQVMQRQALQALNSKSPDKPTPISPG-YQQA 637
 QY 172 DAGNNGSFLNNOEIGWPSRTLDTRVWRPVSPPS-----WQIN-----PNMVF 216
 DB 638 YPMNN---IKQELIQIVQVSL---TQSPDSSPSFLAALGVNASTGCVIATPMNAGT 689
 QY 217 QGVGVGRTSPSPETIMDTLKYSGIGSDNCALSLNSPHOPHDNNNNNNNNNT 276
 DB 690 GSGGCGGLNPPSS-----VGNG-----NSNNGSSGNNNNNST 721
 QY 277 WRASSGFG 284
 DB 722 GNGTSGGG 729
 RESULT 5
 COL6_ARATH STANDARD; PRT; 331 AA.
 ID COL6_ARATH AC Q91Q27;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DE 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Zinc finger protein conserved-like 6.
 GN At1G75540 OR F10A5.24.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicot; Rosidae;
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
 OC NCBI_TaxID=3702;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=cv. Columbia;
 RX MEDLINE=21016719; PubMed=11130712;
 RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
 RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
 RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,

RA Dunn P., Etgu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,
 RA Gall J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Hutzar L.,
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
 RA Kim C.J., Koo H.L., Kremetsekaia I., Kurtz D.B., Kwan A., Lam B.,
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
 RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
 RA Miltchev J., Miranda M., Nguyen M., Nieman W.C., Osborne B.I.,
 RA Pat G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
 RA Sun H., Tallon L.J., Tambunga G., Tortum M.J., Town C.D., Walker M.,
 RA Utechtack T., Van Aken S., Vayberg M., Vysotskaia V.S., Walker M.,
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
 thaliana.";
 RL Nature 408:816-820(2000).
 CC -1- SUBCELLULAR LOCATION: Nuclear (potential).
 CC -1- SIMILARITY: BELONGS TO THE CONSTANS FAMILY.
 CC -1- SIMILARITY: CONTAINS 2 B BOX-TYPE ZINC FINGERS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL, AC006434; AAF87126.1; -
 DR InterPro; IPR000315; Znf_Box.
 DR InterPro; IPR002926; Znf_Constans.
 DR Pfam; PF00643; Zf-B_box; 2.
 DR ProDom; PD007661; Znf_constans; 2.
 DR SMART; SM00336; BBOX; 2.
 DR PROSITE; PS50119; ZF_BOX; 2.
 DR Zinc-finger; Nuclear_protein; Repeat; Multigene family.
 KM ZN_FING 5 47 B BOX-TYPE 1.
 FT ZN_FING 60 102 B BOX-TYPE 2.
 FT DOMAIN 118 126 POLY-SER.
 FT DOMAIN 228 238 POLY-ASN.
 SQ SEQUENCE 331 AA; 36633 MW; 6C2C5A4707976DEA CRC64;

Query Match 7.7%; Score 157.5; DB 1; Length 331;
 Best Local Similarity 25.2%; Pred. No. 0.00018;
 Matches 72; Conservative 39; Mismatches 92; Indels 83; Gaps 15;

QY 114 FCQQ-----CSRFHLPREDEKESCRRLAGHNRPRKQ-----PASLSVLASRYKRI 163
 DB 71 FCQQDRAILC-----KQDSSIHANHTKGRHFLITGVKLSATSSVTKPT 117
 QY 164 APSLYENGDAWGNSFLGNOEIGWPSRSLDTRVWRAPVSSPQWQINPMNVYFSGSGV 222
 DB 118 SKS-----SSSSSSNQDPSVSGSSISNPPLKPLSNAP-OSNNKIQPSKINGDA 167
 QY 223 -----GRTSFSPSEIMDTK-----LES-----YKIGDSCALSLSNPHOPHDNN 263
 DB 168 SVNQWGSTSTISEYMDLPGWHEVDFLDSISLPTYGFSKSGDDCVLPFM-----SPEDDN 223
 QY 264 N-----NNNNNNNNNNNTWR-ASSGFGPMVTYMAQPPAPSOHQYLYLNPWFVKNDNDMPV 319
 DB 224 NTKRNNNNNNNNNNNTVSLPSKNLGIWVQIPLQTLPSYVNYFSG-----QNNN----- 272

QY 320 INIAGYTEPDNCOISSGTAMGEFELSDHHQSRQVQMDENTRAVD 365
 DB 273 IQFGWYMKETSPREVVSFAI-----QNMKQGGN-NKRWID 307

RESULT 6
 KYK1 DICDI STANDARD; PRT; 1584 AA.
 AC P18160;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Non-receptor tyrosine kinase spore lysis A (EC 2.7.1.112) (Tyrosine-
 DE protein kinase 1).
 GN PYXA OR SPLA OR DPYK1.
 OS Dictyostelium discoideum (Slime mold).
 CC Eukaryota; Eukaryota; Dictyostelidae; Dictyostelium.
 CC NCBI TaxID=44689;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=JH10;
 RX MEDLINE=97053827; PubMed=8898241;
 RA Nuckolls G.H., Osherov N., Loomis W.F., Spudich J.A.;
 RT "The Dictyostelium dual-specificity kinase spla is essential for
 RT spore differentiation.";
 RL Development 122:3295-3305(1996).
 RN [2]
 RP SEQUENCE OF 1248-1584 FROM N.A.
 RX MEDLINE=90287147; PubMed=1972546;
 RA Tan J.L., Spudich J.A.;
 RT "Developmentally regulated protein-tyrosine kinase genes in
 RT Dictyostelium discoideum.";
 RL Mol. Cell. Biol. 10:3578-3583(1990).
 CC -1- FUNCTION: ESSENTIAL FOR SPORE DIFFERENTIATION.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED THROUGHOUT DEVELOPMENT WITH A PEAK
 CC DURING THE MOUND STAGE OF MORPHOGENESIS.
 CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES.
 CC -1- SIMILARITY: CONTAINS 1 SAM DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL, U32174; ABA41125.1; -
 DR EMBL, M33785; AAA33202.1; -
 DR PIR; A35670; A35670.
 DR DictyDb; DP03010; Pyka.
 DR InterPro; IPR000715; Euk_Pkinase.
 DR InterPro; IPR001660; SAM.
 DR InterPro; IPR003878; SPRY domain.
 DR InterPro; IPR003877; SPRY receptor.
 DR InterPro; IPR004040; STY_Pkinase.
 DR InterPro; IPR001245; Tyr_Pkinase.
 DR Pfam; PF00068; Pkinase; 1.
 DR Pfam; PF00536; SAM; 1.
 DR Pfam; PF00622; SPRY; 3.
 DR ProDom; PD000001; Euk_Pkinase; 1.
 DR SMART; SM00454; SAM; 1.
 DR SMART; SM00449; SPRY; 3.
 DR SMART; SM00221; STYK; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE; PS50105; SAM DOMAIN; 1.
 KM Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation.
 FT DOMAIN 908 972 SAM.
 FT DOMAIN 403 420 POLY-ASN.
 FT DOMAIN 428 435 POLY-THR.
 FT DOMAIN 449 480 POLY-ASN.
 FT DOMAIN 483 491 POLY-ASN.
 FT DOMAIN 494 508 POLY-ASN.
 FT DOMAIN 512 532 POLY-ASN.
 FT DOMAIN 596 600 POLY-ASN.
 FT DOMAIN 808 811 POLY-PHE.
 FT DOMAIN 1026 1029 POLY-SER.
 FT DOMAIN 1195 1210 POLY-ASN.
 FT DOMAIN 1215 1220 POLY-GLN.
 FT DOMAIN 1224 1233 POLY-GLN.
 FT DOMAIN 1266 1274 POLY-PRO.

```

FT DOMAIN 1289 1561 PROTEIN KINASE.
FT NP BIND 1295 1303 ATP (BY SIMILARITY).
FT BINDING 1316 1316 ATP (BY SIMILARITY).
FT ACT SITE 1417 1417 BY SIMILARITY.
FT CONFLICT 1248 1248 D -> R (IN REF. 2).
FT CONFLICT 1435 1435 V -> L (IN REF. 2).
SQ SEQUENCE 1584 AA; 174304 MW; 5D158945B8E01E3 CRC64;

Query Match 7.5%; Score 152; DB 1; Length 1584;
Best Local Similarity 20.3%; Pred. No. 0.0026;
Matches 97; Conservative 54; Mismatches 133; Indels 174; Gaps 19;

QY 4 GSNQSGPHGPGQASGSGSTSSSPGGLMFG-----QKIYEDGG----- 44
DB 107 GINNGRNVYGP-----SFSSGDVVGCGYDSSSKTYLFTKNGVYLGAQAQVNL 153
QY 45 -----GSSSGGSRNRNVGSGSG--GSGQIPRCVYE 77
DB 154 IGLYPTVGLQNPSSVYINFPGPSYRGAPKPSKQSTIKDSSGSIIPSEDLIPKEEFE 213
QY 78 GCGMDLTNAKGYRRHRCVHSHKT-----PKVTYAGIE-----ORFQCGSRPHQL 124
DB 214 VCRW--SEKQYHGHKHYV--VRNRTAFPLDSPKDTIGCVATQPRGSGFCYFEVYIIDL 269
QY 125 PEPDLKRSRRLAGHNERRRKQDPASLVASRYGRIASLYENGDAGMNGSFLGNQE 184
DB 270 DKQL-----SIGLANLEY-----PTFY-----H 288
QY 185 IGV-PSS---RTLTRVRRRPPVSSPSQINP-----MNVPSGSGVGGRTSPSPSEI 232
DB 289 VGMVPRSYGYHNDGQRKR-----WREPGVNEGSEYSGSYKKGDIIGGLSFTSREI 341
QY 233 MDLTLESYKYGIDGNSC-----ALSLNPHQ----- 258
DB 342 FETKNGMTLGAFFSNVGVFPYSVAFNEPGISITGVFPFPKFSQVITMLKQNVNSTSLV 401
QY 259 .PHDNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 314
DB 402 PNGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 461
QY 315 DMSVYLVAGRTPEPNCQISGTAMGEFELSDHHQSRQYMEDENTRAYSSSHHTN 372
DB 462 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 519

RESULT 7
PK2_DICDI STANDARD; PRT; 1858 AA.
AC P54674;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Phosphatidylinositol 3-kinase 2 (EC 2.7.1.137) (PI3-kinase)
DE (Ptdins-3-kinase) (PI3K).
DE PI3K OR PI3K.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_Taxid=44689;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=AX3;
RX MEDLINE=96009592; PubMed=7565716;
RA Zhou K., Takegawa K., Emr S.D., Pirel R.A.;
RT "A phosphatidylinositol (PI) kinase gene family in Dictyostelium
RT discoideum: biological roles of putative mammalian p110 and yeast
RT vps34p PI 3-kinase homologs during growth and development.";
RL Mol. Cell. Biol. 15:5645-5656(1995).
CC -1- CATALYTIC ACTIVITY: ATP + 1-phosphatidy1-ID-myo-inositol = ADP +
CC 1-phosphatidy1-ID-myo-inositol 3-phosphate.
CC -1- SIMILARITY: BELONGS TO THE PI3/PI4-KINASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U23477; AA085722.1; -.
DR DictyDb; DD01100; PI3K.
DR InterPro; IPR002420; PI3K_C2.
DR InterPro; IPR000341; PI3K_ras_bind.
DR InterPro; IPR001263; PI3Ka.
DR InterPro; IPR000403; PI3_P14_kinase.
DR Pfam; PF00454; PI3_P14_kinase; 1.
DR Pfam; PF00792; PI3K_C2; 1.
DR Pfam; PF00794; PI3K_C2; 1.
DR SMART; SM00142; PI3K_C2; 1.
DR SMART; SM00144; PI3K_Cbd; 1.
DR SMART; SM00145; PI3Ka; 1.
DR SMART; SM00146; PI3K; 1.
DR PROSITE; PS00915; PI3_4_KINASE_1; 1.
DR PROSITE; PS00916; PI3_4_KINASE_2; 1.
DR PROSITE; PS00916; PI3_4_KINASE_3; 1.
DR TRANSFERASE; Kinase; Multigene family.
FT DOMAIN 34 40 POLY-GLY.
FT DOMAIN 166 172 POLY-SER.
FT DOMAIN 185 226 POLY-ASN.
FT DOMAIN 227 235 POLY-THR.
FT DOMAIN 246 253 POLY-SER.
FT DOMAIN 259 268 POLY-ASN.
FT DOMAIN 294 303 POLY-THR.
FT DOMAIN 361 364 POLY-SER.
FT DOMAIN 369 384 POLY-SER.
FT DOMAIN 425 429 POLY-GLN.
FT DOMAIN 439 444 POLY-THR.
FT DOMAIN 445 454 POLY-SER.
FT DOMAIN 562 570 POLY-GLY.
FT DOMAIN 715 727 POLY-THR.
FT DOMAIN 982 990 POLY-GLN.
FT DOMAIN 1015 1049 POLY-ASN.
FT DOMAIN 1598 1858 PI3K/PI4K.
SQ SEQUENCE 1858 AA; 203945 MW; A6C0333D4CDEA03 CRC64;

Query Match 7.2%; Score 146; DB 1; Length 1858;
Best Local Similarity 20.5%; Pred. No. 0.0082;
Matches 62; Conservative 40; Mismatches 114; Indels 86; Gaps 8;

QY 40 FEDGGGGSSSSSGGRNRRVGGGSGSQIIPRCQV-----EGCGMDLTNAKG 88
DB 28 FSNQNSNGS---GGGGSNLSVNSGCSNNSIRKSTLMTNGLPLPSINDGKELLENSK- 83
QY 89 YSRHRCVGHSHKTPKTYAGIEORFQCGSRPHQLPEPDLKRSRRLAGHNERRRK 148
DB 84 -----PKVV-----ELVTFPHKP----- 97
QY 149 QPASLVASRYGRIASLYEN-----GDAG-----MNGSFLGNQEGWPSRRLTD 195
DB 98 -----LSTTHSVHNEIPPALEKEKEKELINTISNGVTYKMTALEILDSTINTPLARRSG 153
QY 196 RYMRPVPSSPSQWQINPMNVPSOGSGVGGRTSFSPEIMDTKLSYKIGDSNCALSLSN 255
DB 154 SIGSKPICN-----NLSSSSSSSTTATTPTTSNNNNNNNNNNNNNNNNNNNNNN 205
QY 256 PHQHDNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 315
DB 206 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 265
QY 316 MS 317
DB 266 NS 267

RESULT 8

```

KNRL DROME STANDARD; PRT; 647 AA.
 ID KNRL DROME STANDARD; PRT; 647 AA.
 AC P13054; GVP64; 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Knirps-related protein.
 GN KNRL OR NR0A2 OR CG4761.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 OC NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89057149; PubMed=2848202;
 RA Oro A.E., Ong E.S., Margolis J.S., Posakony J.W., McKeown M.,
 RA Evans R.M.,
 RA "The Drosophila gene knirps-related is a member of the
 RT steroid-receptor gene superfamily.",
 RL Nature 336:493-496(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20196006; PubMed=10731132.
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle J.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer V., Chape M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abell J.F., Agbayant A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolashikov S.,
 RA Botkova D., Botchan M.R., Bouck J., Brockstein P., Brothier P.,
 RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Fertler S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glaeser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hooten J.D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
 RA Jajani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Kechum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazotto M., Peltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reiner K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kimms I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svrtkals R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster";
 RL Science 287:2185-2195(2000).
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
 CC NRO SUBFAMILY.
 CC -1- CAUTION: Ref.2 sequence differs from that shown due to erroneous
 CC gene model prediction.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
 CC or send an email to license@1sb-sib.ch).
 CC -----
 CC EMBL: X14153; CA33365.1; -
 CC EMBL: AE003591; AAF51627.1; ALT_SEQ.
 CC PIR: S06450; S06450.
 CC HSSP: P20393; 1A6Y.
 CC TRANSFAC: T02773; -
 CC FLYBASE: FBgn0001333; knrl.
 CC InterPro: IPR001628; Znf_C4steroid.
 CC Pfam: PF00105; zf-C4_1_C4steroid.
 CC PRINTS: PR00047; STROIDFINGER.
 CC ProDom: PD000035; Znf_C4steroid; 1.
 CC SMART: SM00399; Znf_C4_1.
 CC PROSITE: PS00031; NUCLEAR RECEPTOR; 1.
 CC Receptor: Transcription regulation; DNA-binding; Nuclear protein;
 CC Zinc-finger; Developmental protein.
 CC DNA_BIND 14 80 NUCLEAR RECEPTOR-TYPE.
 CC ZN_FING 14 34 C4-TYPE.
 CC ZN_FING 51 75 C4-TYPE.
 CC DOMAIN 121 262 GLY-RICH.
 CC DOMAIN 577 597 ASN-RICH.
 CC SEQUENCE 647 AA; 68322 MW; E466FA081DAACDC8 CRC64;
 Query Match 6.1%; Score 125; DB 1; Length 647;
 Best Local Similarity 21.5%; Pred. No. 0.08;
 Matches 89; Conservative 42; Mismatches 139; Indels 144; Gaps 18;
 QY 1 MEMSNGSPGHPGQABSGGSGSTSSPSGGLMFGQKTYFPDGGGGSGSSGGRNRRV 60
 DB 191 VEYSHKRPVAVASVSSVPSDHSNDSSVEVSVKGNPULHGGKNSGSSGADGSH- 248
 QY 61 RGGSGGSGGQIPRCQVSGCGMDLTN-----AKGY----- 89
 DB 249 SCGGGGGGGGGVTTPGRPPQMRKDLSPFLPPLPGLASMPVMPAPLPPSHLLFGYHPAL 308
 QY 90 YSRHRVCGVSHKTPK-----VTVAGIE-----GRFCQCGRF--HQLPFDLEKRS 133
 DB 309 YSHIQ--GLKPTPEQOQAAVAAVHVLFNSSGAGORFAPGTSPFANHQQHKEBQPA 366
 QY 134 CRRLLAGH-----NEBRKRKQPA-----SLSVL 156
 DB 367 PARSPSTHANNHLLTNGGAADLTFRPYDLAVKSGQSGPPPTTKLPKPSKDYISAL 426
 QY 157 A---SRYGRIAPSYEN-----GDAGMNGSFLGNOEIGW-----PSSRTLPTRV 197
 DB 427 VTPNSESGRREVSKRQNEDEDEARADGIDGAHDEDEEDLVSMTPRPHSPAQGEERTPA 486
 QY 198 MRRPVSPSQINPMNVFSGSVGGRTS--FSSPET-----MDTKLESTYKG 242
 DB 487 GEDPPSPSG--QDNIDIL--SMKTGSSLSKSSSPFIEPTEISSDVEKNDTDDDEDLKV 544
 QY 243 IGDSCALSLSNP-----HOPHDNN-----NNNNNNNNNT 276
 DB 545 TPBEIISVRETADEIEEDHSSTETETAKTSTIENTHNNNSISNNNNNNNNNS 598
 RESULT 9
 E78A DROME STANDARD; PRT; 865 AA.
 ID E78A DROME STANDARD; PRT; 865 AA.
 AC P45447; O18395; O18396; GVP64;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Ecdysone-induced protein 78C (DR-78).
 GN E178C OR NR1E1 OR CG18023.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 OC NCBI_TaxID=7227;
 RN [1]

RP SEQUENCE FROM N.A., ALTERNATIVE SPLICING, FUNCTION, AND INDUCTION.
 RC STRAIN=Canton-S; TISSUE=Pupa; and larva;
 RX MEDLINE=94006562; PubMed=8402914;
 RA Stone B.L., Thummel C.S.;
 RT "The Drosophila 78C early late puff contains E78, an ecdysone-
 inducible gene that encodes a novel member of the nuclear hormone
 receptor superfamily.";
 RT Cell 75:307-320(1993).
 RL [2]
 RN SEQUENCE FROM N.A., FUNCTION, AND ALTERNATIVE SPLICING.
 RP STRAIN=Canton-S; TISSUE=Pupa;
 RX MEDLINE=97032934; PubMed=8878682;
 RA Russell S.R.H., Heimbeck G., Goddard C.M., Carpenter A.T.C.,
 RA Ashburner M.;
 RT "The Drosophila Eip78C gene is not vital but has a role in regulating
 chromosome puffs";
 RT Genetics 144:159-170(1996).
 RL [3]
 RN SEQUENCE FROM N.A. (ISOFORM A).
 RP STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10711132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.B., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers J.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt J.G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bernier B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Brotler P.,
 RA Burris K.C., Busam D.A., Butler H., Cadiieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke S., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin R.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fostler K., Gabrielian A.E., Garcia N.S., Gelbart W.M., Glasser K.,
 RA Glisdek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris K.,
 RA Harits N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegian C.,
 RA Jatalin M., Kalish F., Karpen G.H., Ke Z., Kemison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclib J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler P., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Waasamman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers B.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RT Science 287:2185-2195(2000).
 RL [4]
 RN SEQUENCE OF 321-433 FROM N.A. (ISOFORM A).
 RP MEDLINE=94060116; PubMed=8241283;
 RA Martin-Blanco E., Kornberg T.B.;
 RT "DR-78, a novel Drosophila melanogaster genomic DNA fragment highly
 homologous to the DNA-binding domain of thyroid hormone-retinoid
 acid-vitamin D receptor subfamily.";
 RT Biochim. Biophys. Acta 1216:339-341(1993).
 RL [5]
 RN FUNCTION: INDUCES THE EARLY LATE PUFF 78C WHICH TRIGGERS PUPARIUM
 FORMATION AND DEVELOPMENT.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; A/E78A (shown here) and B/E78B;
 CC may be produced by alternative splicing.
 CC -1- DEVELOPMENTAL STAGE: Isoform A is expressed only in mid-pupal

CC stages, while isoform B is maximally expressed in newly formed
 CC prepupae and immediately following isoform A in mid-pupae.
 CC -1- INDUCTION: Both isoforms require ecdysone for activity. Isoform B
 CC also requires ecdysone-induced proteins for maximal expression.
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
 CC NRI SUBFAMILY.
 CC CAUTION: Ref.2 sequence differs from that shown due to
 CC frameshifts in positions 567 to 625.
 CC -1- CAUTION: Ref.3 sequence differs from that shown due to erroneous
 CC gene model prediction.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U01087; AAA19975.1; -;
 CC EMBL; U01088; AAA19976.2; -;
 CC EMBL; X98881; CAA67384.1; ALT_FRAME.
 CC EMBL; X98882; CAA67385.1; ALT_FRAME.
 CC EMBL; AE003593; AAF51692.1; ALT_SEQ.
 CC EMBL; X73045; CAA51523.1; -;
 CC HSSP; P20393; 1A6Y.
 CC TRANSFAC; T02743; -;
 CC FlyBase; FBgn0004865; Eip78C.
 CC InterPro; IPR000536; Hormone_rec_lig.
 CC InterPro; IPR001723; Steroid_receptor.
 CC InterPro; IPR001628; ZnF_C4steroid.
 CC Pfam; PF00104; hormone_rec; 1.
 CC Pfam; PF00105; zf-C4; 1.
 CC PRINTS; PR00398; STRODOMONER.
 CC PRINTS; PR00047; STRODPFINGER.
 CC ProDom; PD000035; ZnF_C4steroid; 1.
 CC SMART; SM00430; HOL1; 1.
 CC SMART; SM00399; ZnF_C4; 1.
 CC PROSITE; PS00011; NUCLEAR RECEPTOR; 1.
 CC Transcription regulation; Alternative splicing.
 CC Zinc; Zinc-finger; Alternative splicing.
 CC Zinc; Zinc-finger; Alternative splicing.
 CC NUCLEAR RECEPTOR-TYPE.
 CC FT DNA_BIND 367 432
 CC FT ZN_FING 367 387
 CC FT ZN_FING 403 427
 CC FT DOMAIN 64 80
 CC FT DOMAIN 182 188
 CC FT DOMAIN 192 202
 CC FT DOMAIN 240 247
 CC FT DOMAIN 271 279
 CC FT DOMAIN 312 315
 CC FT DOMAIN 321 333
 CC FT DOMAIN 336 339
 CC FT DOMAIN 346 349
 CC FT DOMAIN 354 357
 CC FT DOMAIN 481 486
 CC FT DOMAIN 490 500
 CC FT DOMAIN 546 554
 CC FT VARSPIC 2 475
 CC FT CONFLICT 39 40
 CC FT CONFLICT 75 78
 CC FT CONFLICT 321 331
 CC FT CONFLICT 325 326
 CC FT CONFLICT 430 430
 CC FT CONFLICT 433 433
 CC FT CONFLICT 502 502
 CC FT CONFLICT 515 515
 CC FT CONFLICT 733 733
 CC FT CONFLICT 733 733
 CC SQ SEQUENCE 865 AA; 96024 MW; A52734720AFBA492 CRC64;
 CC -----
 CC Query Match 6.1%; Score 125; DB 1; Length 865;
 CC Best Local Similarity 21.0%; Pred. No. 0.11;
 CC Matches 78; Conservative 43; Mismatches 115; Indels 136; Gaps 15;

RA Lamb J.R., Michaud W.A., Sikorski R.S., Hieter P.A.;
 RT "cdc16p, Cdc23p and Cdc27p form a complex essential for mitosis.";
 RL EMBL J. 13:4321-4328(1994).
 CC -1- FUNCTION: EXECUTES ESSENTIAL MITOTIC FUNCTIONS NEAR THE
 CC METAPHASE/ANAPHASE TRANSITION.
 CC -1- SUBUNIT: CDC16, CDC23 AND CDC27 FORMS A MACROMOLECULAR COMPLEX.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: BELONGS TO THE CDC27/NUC2 FAMILY.
 CC -1- SIMILARITY: CONTAINS 8 TPR REPEATS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X79489; CAAS6022.1; ALT_INIT.
 CC EMBL; Z35845; CAAS4905.1; --
 CC PIR; S45825; S45825.
 CC SGD; S0000180; CDC27.
 CC InterPro; IPR001440; TPR.
 CC DR Pfam; PF000515; TPR; 8.
 CC SMART; SM00028; TPR; 5.
 CC DR Cell division; Cell cycle; Mitosis; Repeat; TPR repeat;
 CC KM Nuclear protein.
 CC FT REPEAT 154 187 TPR 1.
 CC FT REPEAT 472 505 TPR 2.
 CC FT REPEAT 540 573 TPR 3.
 CC FT REPEAT 575 607 TPR 4.
 CC FT REPEAT 608 641 TPR 5.
 CC FT REPEAT 643 675 TPR 6.
 CC FT REPEAT 676 709 TPR 7.
 CC FT REPEAT 711 743 TPR 8.
 CC FT DOMAIN 358 391 ASN-RICH.
 CC FT MUTAGEN 613 613 G->D: IN TEMPERATURE SENSITIVE MUTANT.
 CC SQ SEQUENCE 758 AA; 85436 MW; 8612BA4504327A02 CRC64;
 CC
 CC Query Match 6.1%; Score 123.5; DB 1; Length 758;
 CC Best Local Similarity 20.3%; Pred. No. 0.12;
 CC Matches 47; Conservative 30; Mismatches 76; Indels 79; Gaps 5;
 CC
 CC 141 HNEERRKQPSLSTSLSRYSRGIAPSLYENGSDAGNGSFLNGQELGMPSSRTLDTRVWR 200
 CC Db 218 HNNNAASSPFS--TSLSFEPRSQPSLYSKTKGNNNNNNNNNTLFFSSNS----- 267
 CC QY 201 PVSSPSMQINPMNVFSQGSVGGRTSFSSPEIMDTKLESYKIGD-----SNCAISLL 253
 CC Db 268 PPTSSASSFSIIQHSRQOOQANTSIKTKQNTQTEKXPAINSKTSALPNNISMLV 327
 CC QY 254 S-----NPHQHDNN 284
 CC Db 328 SPSSKQPTISSILAKYNNKLLTTPPSKLLNDRHQQNNNNNNNNNNNNNNNNNNNNNN 379
 CC QY 285 PMVTMAQPPAPSOHQYLNPMVFKDNDMSPVYLNGRTEPDNCOISSG 336
 CC Db 380 -----NN 406
 CC
 CC RESULT 12
 CC SRP_DROME STANDARD; PRT: 779 AA.
 CC AC PS2172;
 CC DT 01-OCT-1996 (Rel. 34, Created)
 CC DT 01-OCT-1996 (Rel. 34, Last sequence update)
 CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
 CC DE Box A-binding factor (ABF) (Serpin protein) (GATA-binding factor-B)
 CC (transcription factor GATA-B) (GATA-B).
 CC GN SRP OR ABF.
 CC OS Drosophila melanogaster (Fruit fly).
 CC OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 CC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;

OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxId=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9424445; PubMed=8187633;
 RA Abel T., Michelson A.M., Maniatis T.;
 RT "A Drosophila GATA family member that binds to Adh regulatory
 RT sequences is expressed in the developing fat body.";
 RL Development 119:623-633(1993).
 CC -1- FUNCTION: MAY FUNCTION AS A TRANSCRIPTIONAL ACTIVATOR PROTEIN AND
 CC MAY PLAY A KEY ROLE IN THE ORGANOGENESIS OF THE FAT BODY. BINDS A
 CC SEQUENCE ELEMENT (5'-(TA)GATA-3') FOUND IN THE LARVAL PROMOTERS
 CC OF ALL KNOWN ALCOHOL DEHYDROGENASE (ADH) GENES; ACTS AS A HOMEOTIC
 CC GENE DOWNSTREAM OF THE TERMINAL GAP GENE HKB TO PROMOTE
 CC MORPHOGENESIS AND DIFFERENTIATION OF ANTERIOR AND POSTERIOR
 CC MIDGUT.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- DEVELOPMENTAL STAGE: INITIALLY OBSERVED IN THE ANALGEN OF
 CC THE ANTERIOR AND POSTERIOR MIDGUT AND THE CEPHALIC MESODERM. IT IS
 CC FOUND IN BOTH THE ENDODERMAL AND MESODERMAL GERM LAYERS AND FOR A
 CC BRIEF PERIOD DURING GASTRULATION IT IS EXPRESSED IN THE
 CC AMNIOEROSA. DURING GERM BAND RETRACTION IT BECOMES RESTRICTED TO
 CC THE FAT BODY.
 CC -1- SIMILARITY: CONTAINS 1 GATA-TYPE ZINC FINGER.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X76217; CAAS3807.1; --
 CC HSSP; P17678; IGAT.
 CC DR FlyBase; FBgn0003507; sfp.
 CC DR InterPro; IPR000679; Znf_GATA.
 CC DR Pfam; PF00320; GATA; 1.
 CC DR PRINTS; PR00619; GATZANFINGER.
 CC DR SMART; SM00401; ZNF_GATA; 1.
 CC DR PROSITE; PS00344; GATA_ZN_FINGER_1; 1.
 CC DR PROSITE; PS50114; GATA_ZN_FINGER_2; 1.
 CC KM Transcription regulation; Activator; DNA-binding; Zinc-finger;
 CC KM Nuclear protein; Developmental protein.
 CC FT ZN FING 319 343 GATA-TYPE.
 CC FT DOMAIN 52 55 POLY-ALA.
 CC FT DOMAIN 62 71 POLY-ALA.
 CC FT DOMAIN 81 86 POLY-GLN.
 CC FT DOMAIN 132 137 POLY-GLY.
 CC FT DOMAIN 206 215 POLY-ALA.
 CC FT DOMAIN 254 257 POLY-GLY.
 CC FT DOMAIN 439 445 POLY-GLN.
 CC FT DOMAIN 482 487 POLY-GLN.
 CC FT DOMAIN 539 545 POLY-ASN.
 CC FT DOMAIN 550 557 POLY-ASN.
 CC FT DOMAIN 578 585 POLY-SER.
 CC FT DOMAIN 602 614 POLY-ALA.
 CC FT DOMAIN 700 714 POLY-GLN.
 CC FT DOMAIN 739 744 POLY-GLN.
 CC SQ SEQUENCE 779 AA; 82109 MW; 7D4AD12A241E986D CRC64;
 CC
 CC Query Match 6.1%; Score 123.5; DB 1; Length 779;
 CC Best Local Similarity 21.9%; Pred. No. 0.13;
 CC Matches 75; Conservative 41; Mismatches 122; Indels 105; Gaps 16;
 CC
 CC 16 AESGSGSTESSFSGLMFQKITYFEDGGGSGSSGSGSNRRVRCG-----GS 65
 CC QY 239 AESGSDPYKPSFVWG-----GCGRSKANTSGAASVSCPSNATSAVAS 287
 CC Db 66 GSGQIRCOVEGCGMLDTNAKGYSRHV---CG-----VSKTPKTVAGIBRPFQOC 118
 CC Db 288 GTAATAATTTIDE--HVSRAISRRLSASKRAGLSCNCHTTHTSIMRNPAG--EPVCNAC 343

```

QY 119 S---RFHQLPE-FDLEKSCRRRLAGHNERRRKPPQASISVLASRYGRIAPSLYNGDAG 174
DB 344 GLYKLSHVPRLTYMKKDTIQK-----KKRKKPKGTSEKSKSKXALNAMESGLV 396
QY 175 MNGSLG-----NOEIGMPSRRLDTRVMRPPVSSPMQI-----NMNVFSGSGVGG 222
DB 397 TNCNVAVGLVSSQMDVNDKMPQLDKRYNSYSQPOQQLPOYQQOQQLVADHSSAA 456
QY 223 -----GRTSFSSPEIMDTKLESY-----KGIQDSNC 248
DB 457 SSPHMGSTSL-SPSAMSQHOTHFHQQOQOQOQCGWTCRPTQTCKRSTGSSISSNR 515
QY 249 A-----LSLISNPHQPHDNNNNNN-----NNNNNNNT 276
DB 516 AACSTHPAPHLQHPSPTHQLHNNNNNNSSSLFNNNNNNNNS 558

RESULT 13
MNB_DROME STANDARD; PRT; 843 AA.
ID MNB_DROME STANDARD; PRT; 843 AA.
AC P49657;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Serine/threonine protein kinase minibrain (EC 2.7.1.-).
GN MNB.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berlin;
RX MEDLINE=95161060; PubMed=7857639;
RA Tejedor F., Zhu X.R., Kaltenbach E., Ackermann A., Baumann A.,
RA Canal I., Heisenberg M., Fischbach K.F., Pongs O.;
RT "Minibrain: a new protein kinase family involved in postembryonic
RT neurogenesis in Drosophila.";
RL Neuron 14:287-301(1995).
CC -1- FUNCTION: ROLE IN THE SPECIFIC CONTROL OF PROPER PROLIFERATION OF
CC OPTIC LOBES NEURONAL PROGENY. MUTANTS ARE CHARACTERIZED BY A
CC SPECIFIC AND MARKED SIZE REDUCTION OF THE OPTIC LOBES AND CENTRAL
CC BRAIN HEMISPHERES BUT NO MAJOR ALTERATION IN NEURONAL ARCHITECTURE
CC CAN BE FOUND.
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS: A (SHOWN HERE), B AND C; ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: IN VENTRAL NERVE CORD AND SUPRAESOPHAGEAL
CC GANGLION OF EMBRYOS. IS MOST PROMINENT IN THE MUSHROOM BODY
CC NEUROFIL AND THE OUTER PROLIFERATION CENTER OF THE OPTIC LOBES IN
CC THIRD INSTAR LARVAE.
CC -1- DEVELOPMENTAL STAGE: THE FORMS A AND C ARE PRESENT MAINLY IN
CC EMBRYOS AND PUPAE. BY CONTRAST, FORM B APPEARS TO BE EXPRESSED
CC MOST MARKEDLY IN THIRD INSTAR LARVAE AND PUPAE.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MNB/DYRK SUBFAMILY.
CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-16 IS THE INITIATOR.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X70794; CAA50065.1; -
CC EMBL: X70798; CAA50068.1; -
CC EMBL: X70799; CAA50069.1; -
CC HSRP; P24941; IBUH.
CC FLYBase; FBgn0002777; mnb.
CC FlyBase; FBgn000719; Euk_pkinase.

```

```

DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; Pkinase; 1.
DR Prodom; PD000001; Euk_pkinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR Transferase; Serine/threonine-protein kinase; ATP-binding;
DR Nuclear protein; Alternative splicing; Developmental protein;
KW Neurogenesis.
FT DOMAIN 5 11 POLY-SER.
FT DOMAIN 56 74 BIPARTITE NUCLEAR LOCALIZATION SIGNAL.
FT DOMAIN 99 419 PROTEIN KINASE.
FT NP_BIND 105 113 ATP (BY SIMILARITY).
FT BINDING 128 128 ATP (BY SIMILARITY).
FT ACT_SITE 227 227 BY SIMILARITY.
FT DOMAIN 426 843 GLY/ALA/SER-RICH.
FT DOMAIN 471 474 POLY-SER.
FT DOMAIN 471 474 POLY-SER.
FT DOMAIN 505 508 POLY-SER.
FT DOMAIN 642 646 POLY-SER.
FT DOMAIN 642 646 POLY-SER.
FT DOMAIN 677 680 POLY-SER.
FT DOMAIN 682 692 POLY-SER.
FT DOMAIN 767 776 POLY-SER.
FT DOMAIN 825 828 POLY-SER.
FT VARSPPLIC 536 539 GLM -> DRR (IN ISOFORM B).
FT VARSPPLIC 540 843 MISSING (IN ISOFORM B).
FT VARSPPLIC 536 542 GLMHSV -> VRIARI (IN ISOFORM C).
FT VARSPPLIC 543 843 MISSING (IN ISOFORM C).
FT VARIANT 126 126 A -> T (IN MNB1; REDUCED BRAIN VOLUME).
SQ SEQUENCE 843 AA; 89093 MW; 4AD4EB02CDAED70E CRC64;

Query Match 6.1%; Score 123.5; DB 1; Length 843;
Best Local Similarity 19.7%; Pred. No. 0.14;
Matches 67; Conservative 21; Mismatches 77; Indels 175; Gaps 12;

QY 43 GCGGSGSSGCGRRRRRGSGSGSGGQIRPCQYEGCGMDLTNKGKGYRRHVGYSKT 102
DB 439 GAGGSGSGAGGSGGGG-GGGLAS-----NSSGAVSSSAAAPTAT 482
QY 103 PKVTVAIGIEQFCQCCSRFHQLPEFDLEKSCRRRLAGHNERRRKPPQASISVLASRYGR 162
DB 483 AAATL----- 487
QY 163 IAPSLYNGDAGMNGSLGNOEIGMPSRRLDTRVMRPPVSSPMQINPNVFGSGVGG 222
DB 488 -----AGSSGS-----GSSVGGSSAAQOQAMPPLPLP-LPLPPL-----AGP 526
QY 223 GRTSFSSPEIMDTKLESYKGIQDSNC-----ALSLISNPHQ----- 259
DB 527 G-----GASDQCCHGLMHSVAANAMNFSALISQSNAPPSLANS 568
QY 260 -HDNN-----NNNNNNNNNTWRASGFGPMVTYMA----- 291
DB 569 HHSITSLGSLHISGTSCHNNNSNNNT-RHSRLYGSNMVNMVGHSSGSSNNNS 627
QY 292 -----OPPAPSOHYLNPWPVKDNDNDMP 318
DB 628 ISYPHAMECDPPQMPPPPPNGHGRVPAIMQLQPNYSAP 667

RESULT 14
OMB_DROME STANDARD; PRT; 988 AA.
ID OMB_DROME STANDARD; PRT; 988 AA.
AC Q24432; Q27917; Q9W4K5;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Opsmotor-blind protein (lethal(1)opsmotor-blind) (L(1)omb) (Bifid
DE protein).
GN BI OR OMB OR CG3578.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;

```


OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Oregon-R; TISSUE=Embryo;
RX MEDLINE=92159016; PubMed=1741374;
RA Pfeufferfeld G.O., Roth H., Poock B., Kerschner S., Schwarz H.,
RA Jonschker B., Heisenberg M.;
RT "The lethal(1)optomotor-blind gene of *Drosophila melanogaster* is a
RT major organizer of optic lobe development: Isolation and
RT characterization of the gene.";
RN Proc. Natl. Acad. Sci. U.S.A. 89:1199-1203(1992).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RX MEDLINE=20196006; PubMed=1073132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktoglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Paolis B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.B., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriere S., Fleischmann W.,
RA Foeller C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodde A., Gong P., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwa C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Kerchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Liao X., Lai Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusseken D.R., Paclab J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puti V., Reese M.G.,
RA Rehner K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svrtkvas R., Tector C., Turner R., Venter J., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveril J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Zheng H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195(2000).
RN [3]
RN SEQUENCE OF 1-447 FROM N.A., AND MUTATIONAL ANALYSIS.
RP TISSUE=Larva;
RC MEDLINE=93261414; PubMed=8492800;
RX Poock B., Ballew J., Pfeufferfeld G.O.;
RA "Transcript identification in the optomotor-blind locus of *Drosophila*
RA melanogaster by integrative recombination mapping and PCR-aided
RA sequence analysis of lethal point mutations.";
RL Mol. Gen. Genet. 238:325-332(1993).
RN [4]
RN -1- FUNCTION: ESSENTIAL PROTEIN THAT MAY FUNCTION AS A TRANSCRIPTION
RN REGULATOR. FLIES WITH L(1)OBS MUTATIONS SHOW SEVERE MALDEVELOPMENT
RN OF THE OPTIC LOBES. REDUCTION IN WING SIZE AND AN INCREASED
RN ABDOMINAL PIGMENTATION. THEY DIE DURING THE PUPAL STAGE.
RN -1- SUBCELLULAR LOCATION: Nuclear (Pocentral).
RN -1- TISSUE SPECIFICITY: IN THIRD-INSTAR LARVAE IT IS FOUND IN THE
RN BRAIN REGION THAT WILL DEVELOP INTO OPTIC LOBES AND MORE WEAKLY IN
RN THE THORACIC PART OF THE VENTRAL GANGLION.
RN -1- DEVELOPMENTAL STAGE: THE PEAK PERIODS OF EXPRESSION ARE: MID-
RN EMBRYOGENESIS, THE SECOND DAY OF PUPAL DEVELOPMENT AND IN THE

```
CC ADULT.
CC -I- SIMILARITY: CONTAINS 1 T-BOX DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/ebis/sequence/)
CC or send an email to license@ebi.ac.uk.
CC -----
DR EMBL; M81796; AAA28736.1; -.
DR EMBL; AE003431; AAB26697.1; -.
DR EMBL; S61732; AAB26697.1; JOINED.
DR EMBL; S61729; AAB26697.1; JOINED.
DR EMBL; S61744; AAB26699.1; -.
DR EMBL; S61743; AAB26699.1; JOINED.
DR EMBL; S61955; AAB26699.1; JOINED.
HSP; P24781; 1XBR.
DR FLYBase; FBgn000179; bl.
DR InterPro; IPR001699; TF-T-box.
DR Pfam; PF00907; T-box; 1.
DR PRINTS; PR00937; TBOX.
DR SMART; SM00425; TBOX; 1.
DR PROSITE; PS01283; TBOX_1; 1.
DR PROSITE; PS01284; TBOX_2; 1.
DR PROSITE; PS50252; TBOX_3; 1.
KW DNA-binding; Nuclear protein; Transcription regulation.
FT FT DOMAIN 51 140 ASN-RICH.
FT FT DOMAIN 104 107 POLY-THR.
FT FT DOMAIN 179 184 POLY-SER.
FT FT DOMAIN 229 236 POLY-GLN.
FT FT DOMAIN 238 244 POLY-PRO.
FT FT DNA_BIND 332 513 T-BOX.
FT FT DOMAIN 574 577 POLY-ASP.
FT FT DOMAIN 607 692 ALA-RICH.
FT FT DOMAIN 823 831 POLY-GLY.
FT FT DOMAIN 910 916 POLY-ALA.
FT FT DOMAIN 926 966 GLN/HIS-RICH.
FT FT CONFLICT 10 10 F->L (IN REF. 1 AND 3).
FT FT CONFLICT 216 216 A->P (IN REF. 1 AND 3).
FT FT CONFLICT 511 511 F->L (IN REF. 1).
FT FT CONFLICT 823 823 MISSING (IN REF. 1).
FT FT CONFLICT 976 988 MISSING (IN REF. 1).
SQ SEQUENCE 988 AA; 103992 MW; 032B7A4471743FC9 CRC64;

Query Match 6.0%; Score 123; DB 1; Length 988;
Beet Local Similarity 23.3%; Pred. No. 0.18;
Matches 48; Conservative 33; Mismatches 69; Indels 56; Gaps 9;

QY 173 AGMNGSFLLGNOELGWPSRSLDTRVMRRPVSSPWQIQPMNVVFSQGVGG-----RTSF 227
Db 48 AGSNNNNSGNNTSGNNNSNN-----NTNSTNTNTNLVLAVSPFGGAQLSPGSNH 99
QY 228 SSPEIMDTKLESYKIGIDSCALSLSNPHOPHDNNNNNNNNNNNNNTWTWASSGFGEPT 287
Db 100 SSSMTTTT-----SWTNSSSSNNNNNNSTHNNNNNHTNNNNNNNNNTSQKQH--H 148
QY 288 VTMAQPPAPBSQHQTLNPWPWFKONDNDMSFYLNLGRITPEDNCQISSTGTAMGEFLSDH 347
Db 149 LSTTEEPSPAG--TPPRTIV-----GLPPI-----PPNNNSSSSSSNNNSAAAH 193
QY 348 --HHQSRRQYMEDENTRAYDSSSHHT 371
Db 194 PSNH-----PTAAHS 204

RESULT 15
YK76 YEAST
ID_YK76 YEAST STANDARD; PRT; 1195 AA.
AC P36168;
JT 01-JUN-1994 (Rel. 29, Created)
```

DT 01-JUN-1994 (Rel. 29, last sequence update)
 DT 01-FEB-1995 (Rel. 31, last annotation update)
 DE Hypothetical 137.5 kDa protein in MPL1-PPC1 intergenic region.
 GN YKR096W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 CC NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE OF 1-843 FROM N.A.
 RA Baladron V., Ballesta J.P.G., Bou G., del Rey F., Esteban P.F.,
 RA Garcia-Cantalejo J.M., Garcia-Ramirez J.U., Gonzalez A., Jimenez A.,
 RA Revuelta J.L., Santos M.A.;
 RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 737-1195 FROM N.A.
 RA Gallion L., Dujon B.;
 RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
 CC -1 SIMILARITY: STRONG, TO YEAST YIL151C, ALSO TO THE N-TERMINAL OF
 CC YEAST ALDEHYDE DEHYDROGENASE, MITOCHONDRIAL 1 (ALD1).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; Z28321; CAA82176.1; -.
 DR PIR; S38174; S38174.
 DR SGD; S0001804; YKR096W.
 KW Hypothetical protein: ATP-binding.
 FT DOMAIN 210 250 ASN-RICH.
 FT NP_BIND 465 472 ATP (POTENTIAL).
 SQ SEQUENCE 1195 AA; 137490 MW; D7251DED6B523622 CRC64;
 Query Match 6.0%; Score 122.5; DB 1; Length 1195;
 Best Local Similarity 22.4%; Pred. No. 0.24; Indels 119; Gaps 13;
 Matches 65; Conservative 31; Mismatches 75;
 QY 82 DITNAKGYSHRVGVGHSKTPKVTAVAGIEQRFCCQCSRFHQLPEFDLEKSCRRRLAGH 141
 DB 71 DINSSKRRISSRPFSIDIEGK-----NNDHTYPR-----RTYKE 104
 QY 142 NERRRRKPPASISVLAAR-----YGRAPSLYENGDAWMGSPFLNGEIGWPSGR 191
 DB 105 SEKNPSPRVSSSKRLKRENSVGITQSSALISKSPSENG-----GSTAHEK--W---- 152
 QY 192 TLDTVRMRPVSSPSWQINPMNVFSQGS-----VGG-GRTSSPSPEIMD 234
 DB 153 -----SPENMKRPLNV-SGNSLAFVDAGSDEGSKSEIVGGFQRKSNNSGEIND 199
 QY 235 ---TKLESYKIGIDSNCLSLISNPHOPHNNNNNNNNNNNNNTWRASSGFGPWTVTM 290
 DB 200 KONSARDQDFNNSGNNN-----NNNNHSGNNNDNNNNNNNDNNNNNSNS----- 244
 QY 291 AQPPAPSGHQVLNPPWVFKDNDNMSPLVNLGRYTEPDNCQISGTAWG 340
 DB 245 -----RDNNNNND--DSNREENDSCKPPASVYKRS 272

Search completed: April 10, 2003, 13:53:28
 Job time : 18 secs